Stochastic Switching Dynamics

Simonsen, Maria

DOI (link to publication from Publisher):
10.5278/vbn.phd.tech.00013

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

STOCHASTIC SWITCHING DYNAMICS

BY
MARIA SIMONSEN

DISSERTATION SUBMITTED 2017

AALBORG UNIVERSITY
DENMARK

Stochastic Switching Dynamics

PhD Dissertation

Maria Simonsen

Dissertation submitted May 24, 2017

Dissertation submitted: May 24, 2017

PhD supervisor: Associate Professor Henrik Schiøler
Aalborg University

Assistant PhD supervisor: Associate Professor John-Josef Leth
Aalborg University

PhD committee: Associate Professor Torben Knudsen (chairman)
Aalborg University, Denmark

Research Fellow Manuela Bujorianu
University of Strathclyde, United Kingdom

Professor Henrik Madsen
Technical University of Denmark, Denmark

PhD Series: Technical Faculty of IT and Design, Aalborg University

ISSN (online): 2446-1628
ISBN (online): 978-87-7112-970-0

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2017

Abstract

This thesis treats stochastic systems with switching dynamics. Models with these characteristics are studied from several perspectives. Initially, in a simple framework given in the form of stochastic differential equations (SDEs) and, later, in an extended form which fits into the framework of sliding mode control. It is investigated how to understand and interpret solutions to models of switched systems, which are exposed to discontinuous dynamics and uncertainties (primarily) in the form of white noise. The goal is to gain knowledge about the performance of the system by interpreting the solution and/or its probabilistic properties.

One of the contributions is a convergence result for the Euler-Maruyama method (a numerical scheme which produces realizations of stochastic processes) which indicates that this method is suitable for construction of numerical solutions to stochastic systems with discontinuous dynamics.

The other contributions are of more probabilistic character. Stationary density functions are determined for different models of switched systems and a statistical expression for the stationary variance of the control error for a (mechanical) system designed with sliding mode control is developed.

Synopsis

Denne afhandling behandler stokastiske systemer med skifte dynamik. Modeller med denne karakteristika studeres fra flere perspektiver. I første omgang i en simpel konstruktion i form af stokastiske differentialligninger (SDE'er) og senere i en udvidet version, der kan opstå ved sliding mode regulering. Det undersøges, hvordan man forstår og fortolker løsninger til disse modeller af skifte systemer, når systemer har en diskontinuerlig dynamik samt usikkerhed (primært) i form af hvid støj. Målet er at opnå viden om systemets ydeevne ved at fortolke løsninger og/eller deres sandsynlighedsteoretiske egenskaber.

Et af bidragene er et konvergensresultat for Euler-Maruyama-metoden (en numerisk ordning, der frembringer realiseringer af stokastiske processer), hvilket tyder på, at denne metode er velegnet til konstruktion af numeriske løsninger til stokastiske systemer med diskontinuerlig dynamik.

De øvrige bidrag er mere sandsynlighedsteoretiske. Stationære tæthedsfunktioner er bestemt for forskellige modeller af skifte systemer, og der udvikles et statistisk udtryk for den stationære varians af reguleringsfejlen for et (mekanisk) system designet med sliding mode regulering.

Contents

Thesis Details

Thesis Title:	Stochastic Switching Dynamics
PhD Student:	Maria Simonsen
Supervisors:	Associate Professor Henrik Schiøler, Aalborg University
	Associate Professor John-Josef Leth, Aalborg University

The main body of this thesis consist of the following papers.

(A) M. Simonsen, J. Leth, H. Schiøler and H. Cornean, *A Simple Stochastic Differential Equation with Discontinuous Drift*, published in the proceedings of the Third International Workshop on Hybrid Autonomous Systems 2013.

(B) M. Simonsen, H. Schiøler, J. Leth and H. Cornean, *A Convergence Result for the Euler-Maruyama Method for a Simple Stochastic Differential Equation with Discontinuous Drift*, published in the proceedings of the American Control Conference 2014.

(C) M. Simonsen, H. Schiøler and J. Leth, *Investigation of Random Switching Driven by a Poisson Point Process*, published in the proceedings of the 2015 IEEE Multi-Conference on Systems and Control.

(D) M. Simonsen, H. Schiøler and J. Leth, *Investigations of the Switching Dynamics in Sliding Mode Control*, will be published in the proceedings of the 20th World Congress of the International Federation of Automatic Control 2017.

In addition to the main papers, the following publications have also been made.

(1) H. Schiøler, M. Simonsen and J. Leth, *Stochastic Stability of Systems with Semi-Markovian Switching*, published in Automatica 2014.

(2) H. Schiøler, J. Leth, M. Simonsen and A. Khan, *Stochastic stability of diffusions with semi-Markovian switching*, published in the proceedings of the 54th IEEE Conference on Decision and Control 2015.

Preface

This thesis is submitted as a collection of papers in partial fulfilments for the degree of Doctor of Philosophy at the Section of Automation and Control, Department of Electronic Systems, Aalborg University, Denmark. The work has been conducted in the period from September 2012 to May 2017 under the supervision of Associate Professor Henrik Schiøler and Associate Professor John-Josef Leth.

The thesis contains three parts. Part I is an extended summary of the contributions which contains an introduction including motivation and objectives for the PhD project, a methodology, a summary of the contributions together with conclusion and perspectives. Part II contains the contributions which consist of four papers in the area of stochastic systems with switching dynamics. Part III contains the appendix.

For readers interested in obtaining an overview of the content and contributions of the thesis, it is recommended to read section 5 -"Summary on Motivation and Objective", the chapter "Summary of Contributions" and the chapter "Conclusion and Perspectives".

Acknowledgements

The studies in this thesis have been conducted under the supervision of Associate Professor Henrik Schiøler and Associate Professor John-Josef Leth. At the time I initiated this project almost five years ago, I barely knew what a stochastic differential equation was. I own my gratitude to my supervisors, Henrik and John, who gave me the opportunity to study this area in depth and their ongoing questions which enlightened my understanding and interpretation even further. Special thanks to Henrik for his patience when I repeatedly questioned the practical perspectives of the research project and asked for real-world examples of a simple switched system. Also many thanks to John, who provided and recommended (and discussed with me) much of literature on everything from control theory to stochastic systems to the discipline of mathematical analysis.

And not to forget, I am grateful to Henrik and John for showing me, during many meetings, that true researches fight for their believes and are willing to discuss until a common agreement in a scientific context is settled.

During the PhD period I received a travelling scholarship from the foundation "Travelling Scholarship for Mathematicians" which gave me the opportunity to study abroad for two semesters. Thanks to this foundation, my insight into the mathematical disciplines of stochastic calculus and numerical methods was shaped even further.

I also own thanks to Professor John Lygeros and his research group at the Automatic Control Laboratory, ETH Zürich in Switzerland, who I visited in the spring of 2015. It was an enlightening stay with good meetings and discussions and a research environment which invites for further investigations and developments.

The final thanks goes to Dimo, just for being there and supporting me in the moments I needed it most.

Maria Simonsen
Aalborg University, May 24, 2017

Acronyms and List of Symbols

Acronyms

The following acronyms are used throughout the thesis.

OU	Ornstein-Uhlenbeck
PDE	Partial differential equation
SDE	Stochastic differential equation

List of Symbols

The list of symbols counts for part I and III of the thesis. In the second part of the thesis, the symbols are introduced independently in the individual papers.

C^2	a C^2 function is twice continuously differentiable
E	the state space set and an index set
$\mathcal{F}(\cdot)$	a differential inclusion which generates Filippov solutions
$\tilde{\mathcal{F}}(\cdot)$	a differential inclusion which generates Krasovskii solutions
$\{\mathcal{F}_t\}$	a filtration
$\mathbb{N}$	the natural numbers
$\mathbb{R}^n$	the n-dimensional vector space of real numbers
$\mathbb{R}^+$	the positive real numbers inclusive zero
$\mathcal{T}$	an index set which is a subspace of $\mathbb{R}^+$ which represents often time
$\mathbf{W}$	a d-dimensional Wiener process
$\mathbf{X}^\top$	the transposed vector of $\mathbf{X}$
$\lvert\mathbf{X}\rvert$	the Euclidean norm, that is, for $\mathbf{X} \in \mathbb{R}^n$ is $\lvert\mathbf{X}\rvert^2 = \sum_{i=1}^n X_i^2$

Part I

Introduction

Presentations and State of the Art

This chapter presents the motivation and background for studying stochastic switching dynamics and gives a general introduction to related literature. The chapter ends with a summary of the objective of the PhD project.

1 To Model Real-World Phenomena

Dynamic models of real-world phenomena are used to express the behaviour of physical objects evolving in time. A dynamic model consists of a set of parameters or variables together with a set of rules describing the evolution of the physical object. Examples of dynamical systems are found in thermodynamics, mechanics, chemistry, economics and electronics. More specific examples are the behaviour of a moving pendulum, an electric circuit with relay or a mass-spring mechanical system, [1].

The main purpose of modelling is to reduce a physical system to a set of mathematical equations. Having obtained a model of the system, validation is needed to prove that the mathematical model does indeed describe the real-world situation. However, even a properly developed (deterministic) model is always an approximation of reality, so it is necessary to consider the fact that models might be exposed to imprecisions like errors, deviations and, in general, uncertainties. One method to manage this challenge is to include uncertainties as a part of the modelling.

1.1 Uncertainty in Modelling

In the modelling of real-world phenomena it is often necessary to consider and/or to capture different kind of uncertainties generated by randomness in the behaviour of the system. The phrase "randomness" is a property which can be assigned to parameters or variables of the system which is modelled. Roughly speaking, uncertainty can be modelled in a non-deterministic or in a probabilistic framework.

A non-deterministic framework is often expressed by ordinary differential equations, where the value or position of parameters/variables are unknown but with some

predefined constraints. For example, the disturbance in a non-deterministic model might be upper bounded by a constant without being assigned any specific value or an input can be binary without further information available.

A model in a probabilistic framework is a model where (some of) the variables are assigned a probability measure. In the probabilistic framework it is possible to talk about stochastic or statistic methods (and fuzzy models which will not be discussed in this thesis). The term stochastic is used for to systems which behaviour are generated by a probabilistic model and, therefore, cannot be predicted precisely. The term statistic is used when the target is to assign a qualitative measure based on available data or its probability distribution. Sometimes statistic is referred to as applied probability.

In the sequel, more about noise and disturbance is presented.

Noise and disturbance

Most models of electrical systems are prone to noise in one way or another. It can be noise from the electrical circuit, processor noise, measurement noise from sensors or external disturbance due to impact from uncontrolled factors.

Noise is often classified as small changes from the expected observation which can be parameter dependent or completely independent of the current state of the system.

Disturbance of systems can cause a major change in the evolution of the state position or parameters. This kind of changes can be modelled as jumps, which can be done with the Poisson process in a stochastic frame work [2]. A power cut in an electrically driven system is an example of an unexpected significant change which can be modelled with the Poisson process.

In most cases, the noise and disturbance are modelled by introduction of a random variable or a stochastic process in the model of the system. Observations with error contributions caused by many different factors are likely to be considered as, what is called, white noise. Modelling of white noise is done mathematically with the Wiener process, which, roughly speaking, is defined as the integral-process of white noise [3]. (Further information on the Wiener process follows in section 4.)

Modelling error

Besides noise, another challenge in modelling of systems is inaccuracies of the model itself, also called modelling error. This might be due to imprecise estimates or approximation of the system behaviour. Roughly speaking, inaccuracies of models can be classified into two groups,

(1) structured uncertainties and

(2) unstructured uncertainties.

Group (1) corresponds to inaccuracies of the terms actually included in the model, for example the designed parameters are estimates of parameters which have slow time variation. Group (2) corresponds to inaccuracies of the system order, that is, when

a system might be modelled better with another dimension, but the actual design is chosen due to other factors e.g. because of ease of implementation [4].

Regardless of the reasons for undesirable system performance, the way to control the system behaviour is to introduce a controller into the model.

1.2 A Control System

The control of a physical system can be done electrically, for example, by changing the input voltage or current, or mechanically, for example, with a bimetal thermostat or with regulation of a valve. The combination of a plant (a system) and a controller which influences the plant is called a control system, which is defined as following.

Definition 1 (A control system). *A control system is an object that is driven by a number of inputs (external signals) and as a response produces a number of outputs [5].*

The necessary performance of the incorporated controller can be determined by analysing the mathematical representation of the control system and, depending on the system, different control methods can be applied. The process of developing, analysing and validating the control method is called control design. In the design process a controller is incorporated into the model and it is designed to act according to observations, usually based on state space dynamics and/or time constraints. Since the real world is inherently nonlinear, corresponding models and applied controllers often become nonlinear as well. With this follows the need of advanced control techniques such as linear parameter-varying control, sliding mode control, rule-based control, predictive control, etc.

In summary, control theory deals with the challenges of controlling or guiding a system to behave in a desired way. Therefore, control theory operates on a control system which consists of interconnected components designed to achieve a desired purpose. To be controllable, the system needs to have at least one measurable output and at least one measurable input which influences the measured outputs in a desired way. The control structure of using the output to change or modify the input is called feedback control. The purpose of classical feedback control is to suppress disturbances, compensate for model deviation and to stabilize the system performance. That is, to keep the system outputs close to a desired trajectory or an equilibrium such that the model of the system remains valid.

A common control technique is linear control. In the sequel, this technique is presented together with a discussion of the advantages and disadvantages.

1.3 Linear Control Theory

The class of linear systems is mathematically tractable and can be used to model a large variety of systems. A system is linear if it satisfies the principle of superposition and, in this case, the state space description of the system can be done with the

mathematical discipline of linear algebra. That is, the description of a linear (control) system is often given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

where $x(t)$ is the state, $u(t)$ is the control function, $y(t)$ is the output of the system and A, B and C are matrices which define the operation of the system. As this formulation indicates, it seems natural to apply linear control theory when the system is modelled with linear differential equations (i.e. the system without the control part is $\dot{x} = Ax(t)$ and $y(t) = Cx(t)$). This has many advantages. Linear control theory is a well developed area with relatively simple methods to investigate controllability, observability and stability of the control system [6].

The advantages of linear control methods imply that in some modelling of nonlinear systems a linear approximation is adopted to simplify the control procedure. Alternatively, nonlinear systems may be studied only locally, usually in a neighbourhood of an equilibrium point by linearising the system around such a point. However, linear approximation methods rely on the assumption that the operation range is small enough to ensure validation of the linear model. For large operation range, the linear controller is likely to make poor performances or there will be regions of the state space with unstable behaviour [4].

The work in this thesis does not apply linear control methods due to various reasons. In general, all real-world control systems are nonlinear, and there are systems with nonlinearities which are not linearisable. For systems where both linear and nonlinear control methods are applicable, an individual analysis must be carried out to obtain the most reliable method. For example, a linear control method might bring the best performance to a system with specific parameters, but the nonlinear control method might be more robust to parameter-variations and unmodelled disturbance. Additionally, in some cases nonlinear control might be simpler and more intuitive than linear control, and provide a more robust performance. Other reasons to use nonlinear control methods could be cost and performance optimality.

Sometimes it is preferable that the system can be affected by different control dynamics. In this case the model is designed to change between several subsystems or controllers to improve the execution. These systems fall under the category of switched systems [7]. The next section presents switched systems and introduces the relation to the framework of hybrid systems.

2 Switched Systems

Switched systems arise in many systems for automatic control. They can be found in relatively simple systems such as a compressor in a refrigerator, which keeps the temperature stable, in a more advanced system as an automatic gearbox, or in complex systems such as the total control over a flying aircraft. The advantages of switching

control as compared to continuous/linear controllers are found in robustness to disturbances, simplicity of actuator design and, in some cases, savings in actuator's energy consumption.

The class of switched systems includes systems which consist of a combination of both continuous and discrete dynamics. In many models the discrete dynamics is governed by a supervisor, sometimes called a master, which generates a switching signal based on the switching events. A switching event can be state-dependent or time-dependent.

The continuous dynamics is driven by a set of subsystems influenced by different controllers. The discrete switching signal indicates the current active subsystem and, thus, the active controller. A schematic of a switched system is presented in Figure 1, which is adapted from [8]. In this schematic the discrete dynamics consists of n subsystems, here denoted as controllers. Based on the output function $y(t)$, the supervisor decides whether to keep the current controller or to switch to another controller. The selected controller defines the control $u(t)$, which is included as a part of the input signals. In this model, the switching can be both state-dependent and time-dependent.

A general (deterministic) description of a switched system is

$$\dot{\mathbf{x}}(t) \quad = \quad f_{\sigma(t,\mathbf{x}(t))}(t,\mathbf{x}(t))$$

where, for $t \in \mathcal{T} \subset \mathbb{R}^+$, $\mathbf{x} : \mathcal{T} \to \mathbb{R}^n$ is the continuous state and $f_\sigma : \mathcal{T} \times \mathbb{R}^n \to \mathbb{R}^n$ represents the vector field. The function σ is the switching signal, and it is a piecewise constant function $\sigma : \mathcal{T} \times \mathbb{R}^n \to E$, where E is a finite index set. The switching signal has a discontinuous jump at every switching time, which is defined by every switching event. Thus, the switching signal is constant in between two switching events.

In some cases it is preferred that the system undergoes slow switching, which can be conducted by introducing a dwell time. In this case the switching event is restricted to the situation when the switching time between two switching events is greater than or equal to the dwell time. Often dwell time is introduced to ensure stability of switched systems, which is demonstrated by the related work [9, 10]. However, there are systems where fast switching is unavoidable. One fast switching phenomenon is called Zeno behaviour, which is explained in the sequel.

Zeno behaviour

The dynamics of a (control) system generates switching pattern which, in some cases, can result in undesired behaviour of its model. The following is an example of a real-world phenomenon which generates Zeno behaviour.

Example 2.1 (Bouncing Ball)
The (normalized) equation of motion of a bouncing ball is given by

$$\dot{h} = v, \qquad \dot{v} = -1, \quad (h(t_0), v(t_0)) = (h_0, v_0) \tag{1}$$

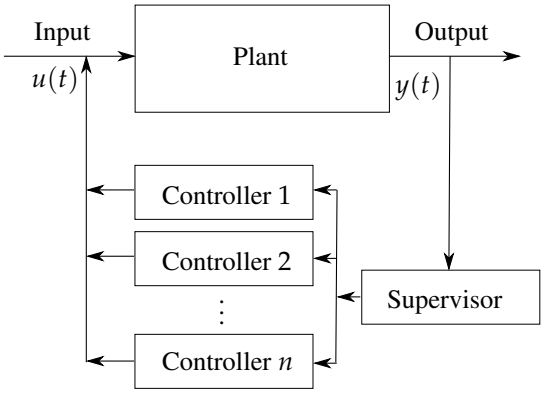

Figure 1: Schematic of a switched system.

where $h(t)$ is the height of a ball above a surface, $v(t)$ is its velocity and (h_0, v_0) are initial conditions. When the ball hits the surface at the switching time t_s, the velocity changes according to the dynamics

$$v(t_s) = -r \lim_{s \to t_s} v(s) \tag{2}$$

where $r \in (0,1)$ is the restitution coefficient. That is, for every switching event the velocity changes sign and decreases in absolute value. As a result, the maximum height above the surface decreases for every subsequent switching event.

Under the initial conditions that for $t = 0$ is $(h(0), v(0)) = (0,1)$, the solution to the equations in (1) is

$$(h(t), v(t)) \quad = \quad (t - \frac{1}{2}t^2, -t + 1)\ ,$$

for $t \in (0, t_{s_1})$ where $t_{s_1} = 2$ is the first switching time. At this time instant the velocity changes according to (2), such that $v(t_{s_1}) = r$. Thus, for $t = t_{s_1}$ is $(h(t_{s_1}), v(t_{s_1}) = (0, r)$. The next switching time t_{s_2} and the corresponding velocity, can also be determined from (1) by resetting the initial conditions. By repeating this procedure at every switching time, the velocity at the n'th switching time t_{s_n} is obtained to be

$$v(t_{s_n}) \quad = \quad r^n$$

and the time between two subsequent switching times is

$$t_{s_{n+1}} - t_{s_n} \quad = \quad 2r^n$$

which implies that the n'th switching time is

$$t_{s_n} \quad = \quad \sum_{i=0}^{n-1} 2r^i\ . \tag{3}$$

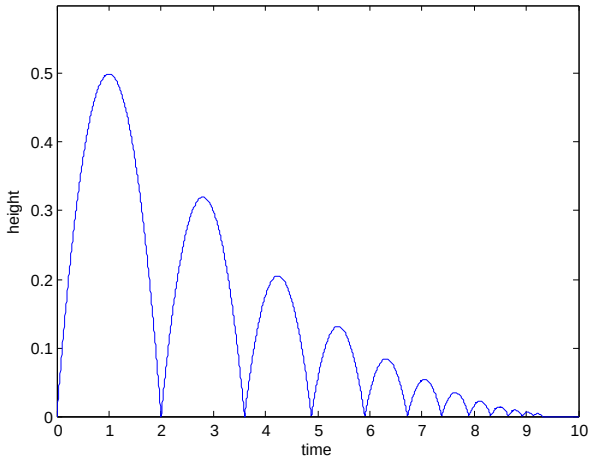

Figure 2: The evolution of the height of a bouncing ball with restitution coefficient $r = 0.8$. The finite accumulation point is a time $t = 10$.

For $n \to \infty$ the sum in (3) is the geometric series which converges and the bouncing ball has a finite accumulation point at time

$$t_{s_\infty} \quad = \quad \sum_{i=0}^{\infty} 2r^i \quad = \quad \frac{2}{1-r} \, .$$

The evolution of the height with respect to time is presented in Figure 2, which also illustrates that the time between two switches decreases as time passes and that the height above the surface decreases. Both situations result in an increasing number of switches in finite time.

As the example illustrates, Zeno behaviour refers to the cases where the switching events happen very fast and very close to each other in terms of state space position, such that the behaviour of the system conducts infinitely many switches in finite time. Eventually, some of these cases will have a finite accumulation point (at least the physical system which is modelled). However, to analyse Zeno executions and to detect potential accumulation points can be complicated and challenging [7]. Therefore, this phenomenon is often excluded from general models available in the literature. This general tendency of ignoring the Zeno behaviour in the design of control systems is one of the motivations behind this study in stochastic switching dynamics.

A Bang-Bang controller

A simple switched system appears in, what is called, Bang-Bang control, on-off control or hysteresis control, which all switch discontinuously between input-values. This

type of controllers can produce fast response, but they are inherently nonlinear. The switching signal is limited to two switching conditions, that is, the system consists of only two continuous subsystems and, correspondingly, two discrete modes. Bang-bang controllers can be found in (closed loop) systems based on an optimization procedure where cost and dynamics depend linearly on the control signal. Other places are in thermostats where the controller regulates the temperature by turning on and off the heater/cooler or in electromechanical relays where an inductor controls a contact mechanism to turn on/off the flow of current [1]. The following is an example of temperature regulation.

Example 2.2 (On-off control)
The evolution of the temperature of a given object (with uniformly distributed temperature) is given by the differential equation

$$\dot{T}(t) \quad = \quad -\frac{1}{c}(T(t) - T_{\text{amb}}) \quad = \quad -\frac{1}{c}\Delta T(t)$$

where c is a positive time constant characteristic of the system and T_{amb} is the ambient temperature. To keep the temperature within the temperature interval $(T_c - \epsilon, T_c + \epsilon)$, a state space dependent controller u_σ is introduced to the system,

$$\dot{T}(t) \quad = \quad -\frac{1}{c}(u_{\sigma(T(t))}(t) + \Delta T(t)) \quad = \quad f_{\sigma(T(t))}(T(t))$$

where the switching signal is

$$\sigma(T(t)) = \begin{cases} 0 & \text{if} \quad T(t) \leq T_c - \epsilon \\ 1 & \text{if} \quad T(t) \geq T_c + \epsilon \\ \lim_{s \nearrow t} \sigma(T(s)) & \text{if} \quad T(t) \in (T_c - \epsilon, T_c + \epsilon) \end{cases}, \qquad (4)$$

with the control defined as $u_0(t) = 0$ and $u_1(t) = \hat{c}T(t)$ where $\hat{c} > 0$ denotes the cooling factor. The evolution of the relation between the discrete and continuous dynamics is presented in Figure 3.

The example above illustrates what is called hysteresis switching. If the temperature was supposed to keep the reference temperature T_c at all times, that is, the case where $\epsilon = 0$, then the switching mechanism is forced to be infinitely fast to compensate for small fluctuations in the temperature. This behaviour is called chattering. Chattering requires high control activity and, therefore, it is undesired in some models of real-world systems because the fast switching in practice might cause problems with the implemented equipment. The fast switching can be avoided with hysteresis switching. The concept is to implement an offset (in example 2.2 denoted ϵ) with respect to the switching time or position (depending on state or time dependent switching) such that the switching events are delayed and the control of the system becomes

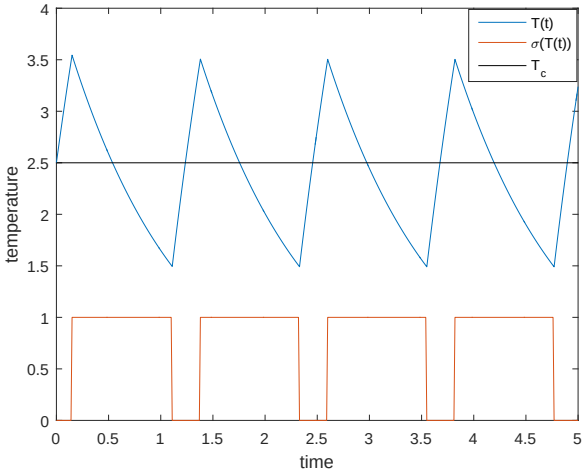

Figure 3: The continuous signal $T(t)$ is controlled by the discrete signal $\sigma(T(t))$ to stay inside the temperature interval $(T_c - \epsilon, T_c + \epsilon)$ with $\epsilon = 1$.

smoother.

A switched system is a special subclass in the class of systems called hybrid systems.

2.1 Hybrid Systems

The class of hybrid systems was introduced in 1966 by H. S. Witsenhausen [11]. Witsenhausen defined hybrid systems as part continuous, part discrete, being described by differential equations combined with multi-stable elements. The paper provides conditions for solutions to hybrid systems to be well-behaved, which means that the solution consists of a finite number of well-behaved arcs and that it can be obtained by admissible control. Furthermore, the paper also states necessary conditions for a class of optimal control problems.

Since the first formal definition, the community has observed that many systems encountered in practice are forms of hybrid systems and that systems with high complexity are suitably modelled within this framework.

There are different models/languages used to describe a hybrid system, which are developed by individuals or research groups based on their individual problems and application aspects. Thus, the literature on hybrid systems is vast, so only a limited amount of contributions are presented here. The common feature of hybrid models is that they present systems which have a discrete and continuous part of the state variable with their own dynamics, but the main challenges in the modelling is to capture the interaction between the discrete and continuous parts.

A common model applied in control engineering is called a hybrid automata. The hybrid automata is a model which combines discrete automata with parameters which

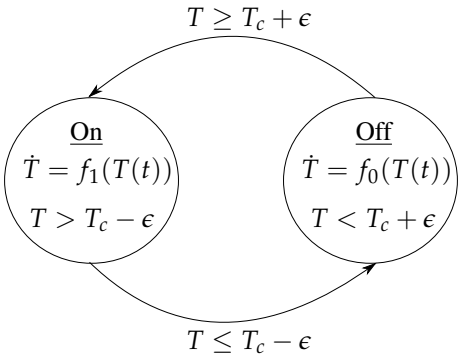

Figure 4: Hybrid automata which regulates the temperature $T(t)$.

evolve continuously. The continuous flow is driven by differential equations and the discrete transitions are determined by an automaton. The hybrid automata captures two kinds of state transitions, discrete jumps which happen instantaneously and continuous flows. A common way to represent a hybrid automata is by a directed connected graph where each node represents a state and each connection between nodes represents a possible transition. A complete description of a hybrid automata can be found in [12].

Example 2.3 (Hybrid Automata)
The example of regulation of temperature presented in example 2.2 can appropriately be modelled as a hybrid automata. The graph representation of the system is given in Figure 4. The graph illustrates that the system switches between state "on" and state "off" when the temperature T leaves the temperature interval $(T_c - \epsilon, T_c + \epsilon)$. Given an initial condition, a solution to the system is characterized by the triple $(\mathcal{T}, \sigma(T(t)), T(t))$, where $\mathcal{T}$ is the index set of discrete switching times, σ is the switching signal given in (4) and $T(t)$ is the trajectory generated according to the continuous dynamics and the switching signal.

In [13, 14] another general hybrid model is presented. This model is based on the mathematical concept of differential inclusion. Both the continuous and the discrete dynamics are defined by differential inclusions called a flow map and a jump map, and belong to a flow set and a jump set, respectively. In [13] several different hybrid models are presented and, in particular, it is illustrated how the hybrid Automata can be reformulated to fit into this hybrid system framework.

In the original formulation of a hybrid system there is no place for uncertainty in the sense of unpredictable and/or probabilistic behaviour. Different authors have made attempts to extend the deterministic hybrid model by introduction of uncertainty into hybrid systems from their own perspectives and application points of view,

i.e. the uncertainty is introduced in different ways. One stochastic hybrid-state model is considered in [15]. In this model, the continuous states are evolving according to stochastic differential equations (SDEs) and the switching dynamics of the individual devices are modelled from the evolution of their corresponding density functions through application of Fokker-Planck equations. In [16], a deterministic hybrid system is governed by a Markov decision process, which makes the system probabilistic. The stochastic contribution is founded in the introduction of policies (sequences of probability measures), which drives the underlying discrete event system. The considered policies are stationary and/or have the Markov property. A similar model is applied in [17], which deals with the class of hybrid piecewise deterministic control systems. Here the discrete state variable is changing with longer idle periods according to a continuous time stochastic jump process, which could be a Markov chain. The model includes also a disturbance parameter which, by assumption, tends to zero. The hybrid model presented in [13] can be reformulated to capture disturbances when they are bounded by a finite norm such that the model becomes non-deterministic. A thorough presentation of hybrid models with different kind of uncertainties included can be found in [12, Chapter 3]. In the following, a general presentation of stochastic hybrid systems is given, which captures a broad representation of uncertainties in real-world applications.

Stochastic hybrid systems

One of the first to introduce uncertainties in hybrid systems by modelling with SDEs was [18]. In this paper the authors replace the deterministic dynamics in the discrete states with SDEs, such that the activation of the different discrete dynamics can occur randomly due to the realization of the stochastic process. A generalization of this model is presented in [19], which claims to include almost all classes of stochastic hybrid processes proposed in the literature at the time of publication.

Definition 2 (A general stochastic hybrid system). *A general stochastic hybrid system is given by H, which is the collection*

$$H \quad = \quad ((Q, d, \mathcal{H}), b, \sigma, \text{Init}, \lambda, R)$$

where

- *Q is a countable/finite set of discrete states.*
- *d is a map given the dimension of the continuous operation modes.*
- *$\mathcal{H}$ maps each $q \in Q$ into an open subset.*
- *b is a vector field (for the continuous dynamics).*
- *σ is a matrix (which defines the impact of the (white) noise).*
- Init *is an initial probability measure.*

- *λ is a transition rate function (which defines the rate of discrete spontaneous jumps).*
- *R is a probabilistic reset kernel.*

[12, Definition 4.1]

A solution to the general stochastic hybrid system is understood as a stochastic process $\mathbf{X}_t = (q_t, \mathbf{Z}_t)$, where $q_t \in Q$ is the discrete process and $\mathbf{Z}_t \in \mathbb{R}^n$ is the continuous process driven by an SDE.

The general stochastic hybrid model is valid under the assumption that the underlying discrete jump process does not admit Zeno behaviour, which implies that the process trajectories do not have chattering or Zeno executions.

Example 2.4 (Stochastic hybrid model of regulation of temperature)
The evolution of temperature in example 2.2 can be exposed to uncertain behaviour. The measured temperature can be influenced by white noise and the control system might experience a power cut.

In the stochastic hybrid model the white noise is modelled by presenting the evolution of temperature by an SDE, while the power cut is modelled as a discrete process with transition rate λ. To model the risk of a power cut, an additional discrete variable $p \in \{0, 1\}$ is introduced, which takes the value $p = 0$ in case of power cut and $p = 1$ when power is available. The stochastic hybrid automata of the system is presented in Figure 5.

The solution to this model is a stochastic process $\mathbf{X}_t = (q_t, T_t), q_t \in Q = \{$On, Off, Power off$\}$ which exhibits forced and spontaneous discrete transitions.

Recent literature on alternative models of stochastic hybrid systems are [20, 21]. The authors to [20] consider a class of stochastic hybrid systems where uncertainty is implemented in an assumption on state-dependent switching. The switching is governed by an intensity function defined in an "ϵ-area" around boundaries between local systems such that the switching between subsystems occurs within this small area of the switching surface. The closer the state trajectory approaches the boundary, the greater is the probability that the state trajectory switches to a new local system. The work [21] considers hybrid systems, which are exposed to time-delays in the switching process. The time-delay can be due to small measurement noise or general delays in the system performance. Therefore, switching events can happen at the boundary or at a time-delayed moment. The study focuses on formulating the generalized concept of solutions into hybrid systems with delay.

One control method where the switching mechanism is a major part of the performance of the controller is in sliding mode control. Following, this control method is presented.

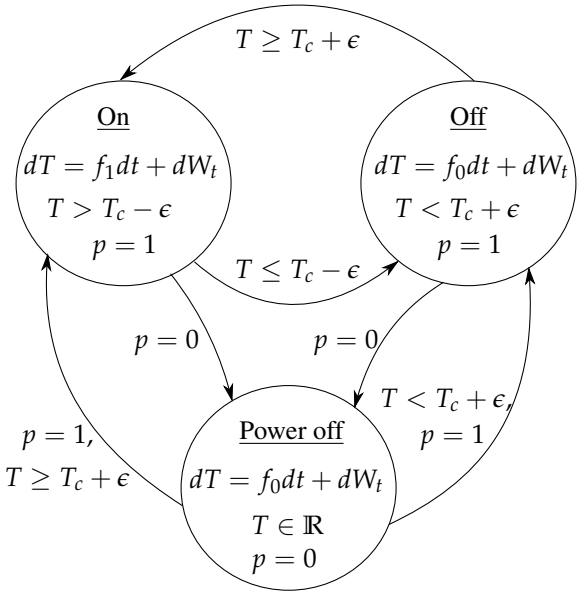

Figure 5: Stochastic hybrid automata for the regulation of temperature presented in example 2.4.

3 Sliding Mode Control

Some nonlinear systems are appropriately controlled with a sliding mode controller. Generally, a sliding mode controller applies a discontinuous control signal to enforce the system state first to approach, and afterwards to slide along a prescribed sub-manifold of the state space. When the system has reached the desired set of operation points, the purpose of the controller is to respond appropriately to any fluctuation such that this position is maintained. Books about application and analysis of sliding mode control are [22, 23].

The design of a sliding mode controller is done on two steps:

(1) Definition of the sliding surface, which is a stable sub-manifold of the state space. The sliding surface is supposed to guide the system towards the desired operation points.

(2) Design of the controller that drives the object/state to the sliding surface.

A well designed sliding mode controller ensures that the sliding mode is reached in finite time and that the controller is sufficiently robust to disturbance.

In the design of a sliding mode controller, the robustness is ensured by knowledge of the disturbance. This "knowledge" could be an upper bound on the maximum disturbance or a measure on the expected disturbance. In most guides on design of

sliding mode controllers the upper bound approach is applied. This means that the sliding mode controller is designed such that robustness is guaranteed under the assumption that the disturbance is finite or bounded by known constraints (i.e. functions, norms or constants) [24].

For systems which are exposed to white noise, only the probabilistic behaviour of the disturbance is known. In these cases it is possible to make a stability analysis of the system with stochastic Lyapunov theory [25]. As special cases, [26] presents stability of stochastic systems with Markovian switching and [27] presents stability results for a linear stochastic system with a sliding mode approach. Even though the stability of stochastic systems can be proven by construction of a stochastic Lyapunov function, this method neither guarantees/ensures the existence of a solution to the stochastic system, nor validates the process of generating numerical solutions.

One of the challenges in the practical implementation of the sliding mode controller is that the switching control signal causes the system to chatter in the neighbourhood of the sliding surface. There are developed methods to reduce this problem. One solution is to replace the hard discontinuous dynamics of the controller with a smooth approximation within a thin boundary layer of the sliding surface [4]. Often this is done by replacing the sgn-function with the saturation-function. Another solution is to implement a hysteresis around the sliding surface such that the switching time gets delayed and the system is prevented from fast switching.

The practical implementation of a smooth control function near the sliding surface is suitable for some systems. However, for some cases this smooth function, often modelled as the saturation function, is inefficient and exposed to loses such as undesirable heat dissipation. In this case a more efficient solution is direct implementation of the switching procedure as fast as possible. Of course this method requires that fast state changing is achievable. This method has successfully been implemented, among other examples, in DC motors and loudspeakers [28].

A smooth solution, which can capture the characteristic system behaviour generated by the state moving along a sliding surface, is denoted a Filippov solution. This concept is defined in the sequel and compared with the Caratheodory solution and the Krasovskii solution.

3.1 Solutions to Systems with Discontinuous Dynamics

Solutions to vector fields with discontinuous dynamics do not exist in the classic understanding, as a continuous differentiable solution to the vector field. Instead, an alternative concept, denoted the Filippov solution, is developed by A. F. Filippov [29]. For continuity points of the vector field, the Filippov solution is identical with a classic solution and in discontinuous points it is constructed to be continuous as well. The idea behind the Filippov solution is to introduce a set of directions which are determined by the vector fields in the neighbourhood of points. The Filippov solution is then chosen among the possible options given by the set of directions.

Let $\mathbf{x} \in \mathbb{R}^n$ be a vector and let $f(t,\mathbf{x})$, $f : \mathcal{T} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a piecewise

continuous vector function. Consider the system driven by the differential equation

$$\dot{\mathbf{x}} = f(t, \mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \ . \tag{5}$$

A classic solution to (5), also called a Caratheodory solution, is an absolute continuous curve that satisfies the integral version of (5), that is,

$$\mathbf{x}(t) \quad = \quad \mathbf{x}(t_0) + \int_{t_0}^{t} f(s, \mathbf{x}) ds, \quad t > t_0 \ .$$

The Caratheodory solution follows the direction of the vector field at all times but the differential equation (5) does not need to be satisfied on a set of measure zero [30].

If the vector field $f(t, \mathbf{x})$ has discontinuous dynamics such that no Caratheodory solution exists, a solution to a redefined model of the system can be considered by constructing a differential inclusion from $f(t, \mathbf{x})$ which can generate a Krasovskii solution or a Filippov solution.

The differential inclusion $\tilde{\mathcal{F}}(t, \mathbf{x})$ is constructed for each point $(t, \mathbf{x})$ in the domain of $f(t, \mathbf{x})$. For continuity points of the vector field, the differential inclusion $\tilde{\mathcal{F}}(t, \mathbf{x})$ is equal to $f(t, \mathbf{x})$. For discontinuity points of the vector field, the differential inclusion is defined as the (closed) convex hull of all possible vector fields in the neighbourhood of the discontinuous point. That is, a function $\mathbf{x}$ is a Krasovskii solution to (5) if

$$\dot{\mathbf{x}} \in \tilde{\mathcal{F}}(t, \mathbf{x})$$

for almost all $t \in \mathcal{T}$, where for each $\xi \in \mathbb{R}^n$ is

$$\tilde{\mathcal{F}}(t, \xi) = \bigcap_{\epsilon > 0} \bar{co} \left\{ f(t, \xi + \bar{B}_\epsilon(\mathbf{0})) \right\}$$

where $\bar{co}$ means closed convex hull and $\bar{B}_\epsilon(\mathbf{0})$ is a closed ball centered at $\mathbf{x} = \mathbf{0}$ of radius ϵ [14].

A Filippov solution to (5) for $t \in \mathcal{T}$ is a solution of the differential inclusion

$$\dot{\mathbf{x}} \in \mathcal{F}(t, \mathbf{x}) \ ,$$

where

$$\mathcal{F}(t, \mathbf{x}) = \lim_{\epsilon \to 0} co \left(\bigcap_{\mu(A)=0} \{ f(t, \mathbf{y}) : \mathbf{y} \in B_\epsilon(\mathbf{x}) \backslash A \} \right)$$

where co means convex hull, $\mu(A)$ is the measure of the set A and $B_\epsilon(\mathbf{x})$ is an open ball centered at $\mathbf{x}$ of radius ϵ. That is, a Filippov solution is an absolutely continuous vector-valued function that satisfies the differential inclusion for almost all $t \in \mathcal{T}$.

Both a Caratheodory solution and a Filippov solution are also a Krasovskii solution but the converse is not true.

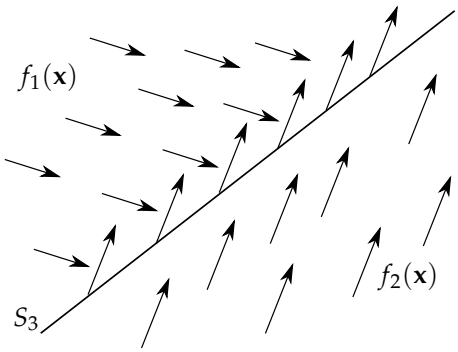

Figure 6: A system with discontinuous dynamics.

As the definition of the differential inclusions indicates, the Filippov solution has similarities with the Krasovskii solution, but they distinguish from each other at the sets of measure zero. Where the Filippov solution excludes sets of measure zero, the Krasovskii solution does not. Therefore, a Krasovskii solution can be significantly different than the Filippov solution. To enlighten the difference further, consider the following example.

Example 3.1 (Different solutions)
Let S be a two-dimensional state space partitioned in three subspaces S_1, S_2 and S_3, where S_3 defines a sliding surface. Let $\mathbf{x}$ be driven by the dynamics

$$\dot{\mathbf{x}} \quad = \quad f_i(\mathbf{x}) \text{ for } \mathbf{x} \in S_i, \quad i \in \{1,2,3\}, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{6}$$

where $f_1(\mathbf{x}), f_2(\mathbf{x}) = f_3(\mathbf{x}) \neq \mathbf{0}$ are constant vector fields. Figure 6 illustrates such a system. Given an initial condition $\mathbf{x}_0$, a Caratheodory solution to the system exists until the time instant where the trajectory hits the sliding surface. At the sliding surface there is no existence of Caratheodory solution due to the discontinuous dynamics. However, both a Filippov solution and a Krasovskii solution exist for all $t \in \mathcal{T}$. A Filippov or Krasovskii solution will follow the vector field and, at the surface S_3, the solution will slide along the surface in the direction generated by the convex hull of the vector fields f_1 and f_2.

Now consider a variation of the system where the constant vector field $f_3(\mathbf{x})$ is redefined to be parallel with the surface S_3. A Filippov solution with initial condition in x_0 is presented in Figure 7. This solution is also a Krasovskii solution. However, this system admits one more Krasovskii solution, which is presented in Figure 8. Notice that this particular Krasovskii solution is also a Caratheodory solution.

Actually, the combination of the two solutions presented in Figure 7 and 8 is also a Krasovskii solution. This means that the Krasovskii solution can slide along

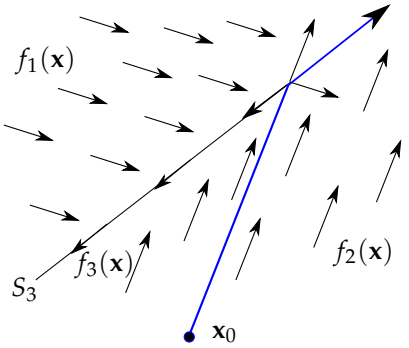

Figure 7: A Filippov and a Krasovskii solution to a system with discontinuous dynamics.

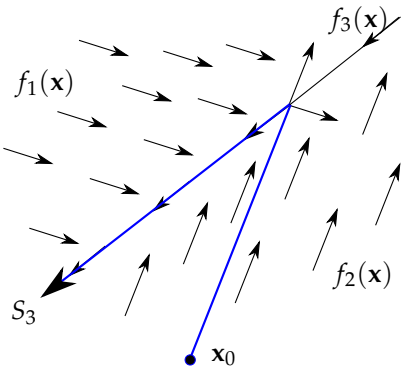

Figure 8: A Caratheodory and a Krasovskii solution to a system with discontinuous dynamics.

the sliding surface and change the sliding direction without violating the dynamics defined by the differential inclusion $\tilde{\mathcal{F}}(t, \mathbf{x})$.

The Filippov solution is often more applicable than the Krasovskii solution, since it is not exposed to odd or unclear behaviour. The solution in the following example is defined as the Filippov solution.

Example 3.2 (A Filippov solution to a version of temperature regulation) Consider a system driven by the dynamics given in example 2.2, but with $\epsilon = 0$, that is,

$$\dot{T} = f_{\sigma(T(t))}(T(t)) = \begin{cases} f_0 & \text{if } T(t) < T_c \\ f_1 & \text{if } T(t) > T_c \end{cases} . \tag{7}$$

A phase plane plot for the system is presented in Figure 9. The point of discontinuity is $T(t) = T_c$. Let the state space be partitioned in subspaces $S_0 = \{T : T < T_c\}$, $S_1 = \{T : T > T_c\}$ and $S = \{T_c\}$. Then a Filippov solution to the switched system (7) is a solution T which satisfies the differential inclusion

$$\dot{T} \in \mathcal{F}(T)$$

where $\mathcal{F}$ is defined as follows

$$\mathcal{F}(T) = \begin{cases} f_i(T) & \text{if } T \in S_i \\ \{\alpha f_0(T) + (1-\alpha) f_1(T) : \alpha \in [0,1]\} & \text{if } T \in S \end{cases} .$$

A Filippov solution with initial temperature $T(0) = 4.5$ is presented in Figure 9.

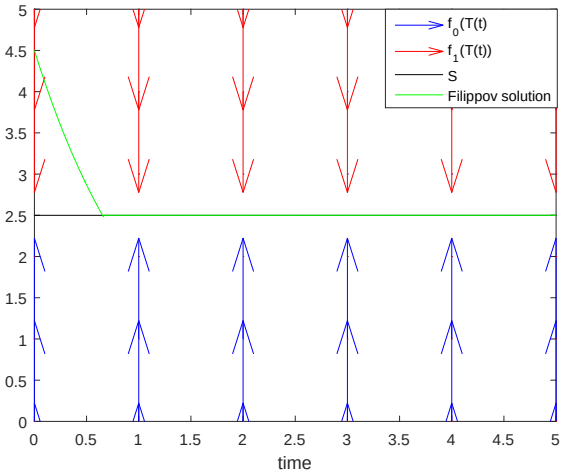

Figure 9: Phase plane plot of the vector fields in system (7) together with a Filippov solution to the system.

In the situation where the Filippov solution slides along the sliding surface, the sliding velocity can be determined. This is done in [31] based on two regularization methods. One of these methods implements a small fixed delay into the deterministic system and deduces the sliding velocity under this restriction. The other method determines the sliding velocity by implementing weak noise such that the considered model is an SDE.

Stochastic differential equations (SDEs) are differential equations with one or more additional terms describing some kind of stochastic influence on the system behaviour. In the next section, modelling with SDEs is discussed.

4 Modelling with SDEs

SDEs are useful frameworks to represent real-world phenomena with uncertainty through mathematical models. This section gives a further introduction to the stochastic behaviour and challenges behind modelling with SDEs.

One of the first to observe and document stochastic behaviour was the Scottish botanist Robert Brown, who noticed the erratic motion of pollen grains suspended in liquid [3]. This random behaviour was named Brownian Motion. The physical behaviour of Brownian Motion can be illustrated with a glass of liquid and a drop of ink. Initially, when a drop of ink is immersed into the liquid, the ink is concentrated on a small area, but over time the motion of the ink-particles resulting from their collision implies that the ink is scattered over the complete area. The concept is illustrated in Figure 10. The mathematical description of Brownian motion is called a Wiener process after Norbert Wiener, who formalized the properties of Brownian motion. A distinctive property of the Wiener process is that almost all realizations are continuous

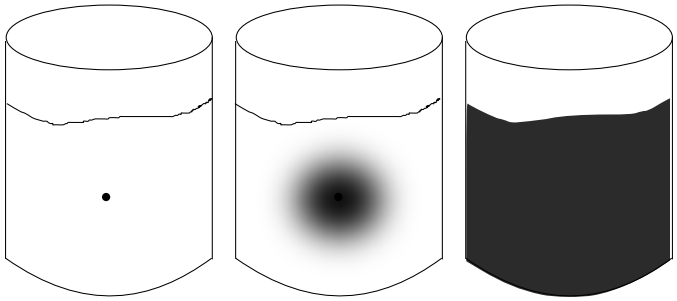

Figure 10: The scattering of a drop of ink in liquid over time illustrates the physical behaviour of Brownian Motion.

but nowhere differentiable. When Brownian motion is described by a Wiener process, the nondifferentiability of the Wiener process indicates that the particles under observation do not possess a velocity at any time [32].

The combination of the stochastic Wiener process and an ordinary differential equation was merged together into stochastic differential equations (SDEs) and the formal mathematical foundation was developed by Itô and Stratonovich [33, 34]. Both Itô and Stratonovich define a stochastic integral over the Wiener process and the corresponding stochastic calculus. There is a difference in the definitions of the two integrals but it is possible to transform one of the representations into the other (see section 6.2). If nothing else mentioned, the stochastic integral applied in this thesis is the Itô integral.

4.1 The Concept of Solutions

One of the motivations to model real-world phenomena with mathematical models is to be able to predict or gain knowledge of the system behaviour without actually observing the physical system in action. A natural interest is, therefore, to understand and to interpret the solutions to the mathematical formulation of the system, that is, to observe the system behaviour based on the stochastic formulation of the system.

An SDE is a combination of a drift coefficient and a diffusion coefficient. Depending on properties of the drift and diffusion coefficient, different results on existence and uniqueness of solutions to the SDE are proven. Generally, if standard regularity conditions as measurability, Lipschitz continuity and boundary conditions are valid for the drift and diffusion coefficient, then a solution to the SDE exists.

The solution to an SDE is a stochastic process, which is defined with respect to a probability space consisting of a set of events, a sigma-algebra and a probability measure (a probability distribution). A solution can be classified as weak or strong. Roughly speaking, a weak solution is a solution in probability while a strong solution makes the SDE hold almost surely with a given initial condition. When a weak so-

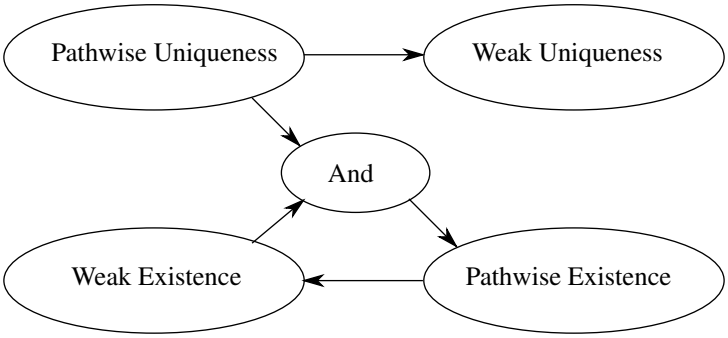

Figure 11: Relation between classifications of solutions. In the discussion of solutions the term pathwise is used interchangeable with strong.

lution exists, it is said that the SDE has weak existence and when the strong solution exists, it is said that the SDE has pathwise existence. A strong solution is also denoted a pathwise solution. It is clear that if an SDE admits a strong solution, then the weak solution exists as well.

In addition to the existence of solutions, uniqueness of solutions is also a question of interest. A solution can be unique both in a weak sense and in a strong sense. A weak solution is often represented as a pair (X_t, W_t), which is assigned a certain probability measure. Then the SDE is said to have weak uniqueness if the joint laws of every two weak solutions, (X_t, W_t) and $(\tilde{X}_t, \tilde{W}_t)$, to the SDE under their respective probability measures are equal [35]. This means that the probability distributions of the solutions are equal. The SDE is said to have strong uniqueness if the probability that any two strong solutions of the SDE are different is zero.

Figure 11 gives an overview of the relation between the different classifications of solutions. The arrows should be understood as "implies", such that whenever the SDE has pathwise uniqueness, it also has weak uniqueness and so forth.

In [36], Yamada and Watanabe state the famous result that the existence of a solution (on some probability space with some Wiener process) and the pathwise uniqueness imply the existence of a strong solution. This is illustrated by the "And" in Figure 11.

A clear and comprehensive exposition of solutions and classifications can be found in [37], which also includes several examples to illustrate the different concept of solutions.

4.2 The Discontinuous Drift Challenges

The representation of a switched system in a stochastic framework often provides discontinuous expressions in the drift coefficient. The existence of solutions to SDEs with discontinuous dynamics is a field which has gained much of attention over the past. Some of the results developed in this area are presented in the following.

In 1974 A. K. Zvonkin presented a paper which transforms an SDE with a drift coefficient to an SDE without a drift coefficient [38]. This can be an advantage, since in certain cases processes which only depend on the diffusion coefficient can be solved easier. Furthermore, it made it possible to construct strong solutions of SDEs with some kind of "bad" drift coefficient. A. J. Veretennikov extended the result from [38] and showed that sufficient conditions for a strong solution are: the diffusion coefficient is equal to the identity matrix and the drift coefficient is a bounded measurable function [39].

To the best of our knowledge, strong existence and uniqueness are established for SDEs with discontinuous bounded drift, but it is not proven yet for SDEs with unbounded discontinuous drift coefficients. However, in implementation of mechanical systems, the main force, which affects the motion of a plant, depends on the actual position of the plant. Thus, the drift coefficient is linear (at least part it) and unbounded. Therefore, there is still need for additional research in the area of SDEs with (unbounded) discontinuous drift coefficients.

In cases where a given SDE does not provide a closed form solution explicitly, it is possible to construct approximated realizations of solutions to the SDE. This may be done with numerical methods.

4.3 Numerical Methods

A numerical method is a mathematical tool designed to solve problems numerically. A good numerical method is inexpensive in terms of computation cost and provides high accuracy. Furthermore, the method has to provide a reliable bound on the error between the numerical solution and the exact solution. Presentation of different numerical methods together with their algorithmic implementation can be found in [40] and [41].

The Euler-Maruyama method is a numerical method which applies to SDEs [42]. It is widely used because of its easy implementation and low computation time together with reasonably good rate of convergence. Another numerical method is the Milstein Scheme [40]. This method is developed under the same principle as the Euler-Maruyama method (Taylor approximation), but it is extended by iterative applications of the Itô formula. This implies that the approximation error decreases and the rate of convergence improves. The payback is longer computation time. Strong Taylor approximations are also used to produce even better numerical methods in terms of accuracy and convergence properties when precision is required.

In general, the Euler-Maruyama method guarantees weak and strong convergence to the stochastic process, if the drift and diffusion coefficient satisfy appropriate regularity conditions. There are many sub-classes of SDEs. Hence, there exists a vast amount of work investigating convergence properties of the Euler-Maruyama method for different sub-classes. Among others, [43] and [44] prove convergence for the class of SDEs with Markovian switching and for SDEs with respect to semimartingales, respectively. The recent years, different results have proved that the Euler-Maruyama

method is also suitable for classes where the SDEs have irregular drift and diffusion coefficient.

I. Gyöngy and N. Krylov have proved that it is possible to construct Euler-Maruyama approximations which converge uniformly on bounded intervals, in probability, to a stochastic process. This process is the strong solution of a specific SDE with discontinuous drift and with a diffusion coefficient which does not satisfy the linear growth condition [45].

Weak convergence of approximated solutions produced with the Euler-Maruyama method is presented in [46]. Here it is assumed that the drift and diffusion coefficient are locally bounded measurable functions and that they are locally continuous over each set of pairwise disjoint sets. That is, the drift and diffusion coefficient are allowed to have discontinuous points.

In [47] a weak rate of convergence of the Euler-Maruyama method is presented when the coefficients are Hölder continuous. In particular, the rate of convergence is determined without assuming the Lipschitz condition. Furthermore, the linear growth condition is assumed for the diffusion coefficient and it is allowed to be discontinuous in a set of points which have Lebesgue measure zero.

A couple of novel results are [48] and [49]. In [48] a transformation technique is presented which is applied to construct a numerical scheme based on the Euler-Maruyama method to prove existence and uniqueness for SDEs with discontinuous drift coefficient. The discontinuity of the drift coefficient is limited to a finite number of jumps and everywhere else it is assumed Lipschitz. In [49] is proven strong convergence of the Euler-Maruyama method when both the drift and diffusion coefficient are allowed to be discontinuous.

5 Summary on Motivation and Objective

One of the challenges which occur in the process of formulating a model of a physical real-world system into a theoretical framework is with respect to the accuracy of the model. It can be challenging to interpret and evaluate the mathematical models such that the physical behaviour is interpreted and understood in the best way possible and that proper statements on the system performance can be given.

In modelling of switched systems there is a general tendency of ignoring the Zeno behaviour because of its complications in the system analysis. Furthermore, the general stochastic hybrid model proposed in [19] and [12] is developed under the assumption that the stochastic process is not subject to Zeno executions. This assumption contradicts with the general behaviour of some switched systems modelled from real-world phenomena.

Another challenge in modelling of switched systems is that the behaviour of the switching mechanism (or control) induces discontinuity into the model. This lack of continuity imposes a significant difficulty as the problem is rendered outside the scope of the classical initial value setting. This introduces the necessity to study irregular

SDEs.

To evaluate a control system designed in the framework of SDEs, stochastic stability analysis is carried out by applying a stochastic Lyapunov function. However, even though a system is proven to be stable in a stochastic sense, this analysis does not guarantee or ensure that any valid solution to the system exists or whether the solution can be constructed numerically.

Based on the discussion above, the objective of this thesis is to study the behaviour of control systems which can be modelled with discontinuous dynamics in a stochastic framework. The initial focus of the study is the simple SDE with discontinuous drift given by

$$dx_t \quad = \quad -k\,\mathrm{sgn}(x_t)dt + dW_t, \quad x_0 = 0 \tag{8}$$

where x_t is the state, $k > 0$ is a (control) system parameter and W_t is the Wiener process. The notion of solution and probabilistic results, numerical results, auto- and covariance results are investigated for systems modelled with this dynamics to improve knowledge on system performance.

A local behaviour of the standard Wiener process W_t is that in every time interval of the form $[0,\epsilon)$, with $\epsilon > 0$, W_t has infinitely many zeros [32]. That is, the process $dx_t = dW_t$ crosses zero infinitely many times and since the drift coefficient $-k\,\mathrm{sgn}(x_t)$ also drives the process x_t towards zero, it seems natural/reasonable to expect that the solution x_t to the SDE in (8) would have characteristics similar to the Wiener process, i.e. x_t might also cross zero infinitely many times in finite time. Therefore, the Zeno phenomenon can not be a property excluded from the framework of the modelling.

Due to [39], it is known that the solution to the SDE in (8) exists. However, even though existence and uniqueness are established for the bounded discontinuous drift coefficient case, the specific characteristics of the obtained solutions, such as transient and stationary distributions as well as auto- and cross-covariance characteristics, seem less studied.

The existence of solutions to SDEs with unbounded discontinuous drift coefficient is still not established in the existing literature even though such models do appear in modelling of (mechanical) systems. Therefore, it is also of interest to investigate solutions and performance properties of such models.

Methodology

This chapter presents theoretical tools which are suitable to investigate, analyse and interpret modelling of dynamical systems. Readers familiar with SDEs and related theories can skip this chapter.

6 Stochastic Evolution

The random evolution of a system can be described by a stochastic process. In the following, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(E, \mathcal{G})$ be a measurable space, where

- Ω is the sample space
- $\mathcal{F}$ is a σ-algebra of subsets of Ω
- $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$
- E is the state space
- $\mathcal{G}$ is a σ-algebra of subsets of E

Then a stochastic process is defined as follows.

Definition 3 (Stochastic process). *A stochastic process is a map $X : \mathcal{T} \times \Omega \to E$ consisting of a collection of random variables.*

The random variables take values in the state space E of the stochastic process. The index set $\mathcal{T}$ often represents time. If $\mathcal{T}$ consists of discrete time instants, X is a discrete time stochastic process and if $\mathcal{T} = [0, \infty)$, X is a continuous time stochastic processes. Thus, a stochastic process $X = X(t, \omega)$ is a function of two variables. The process $X(t) = X(t, \cdot) : \Omega \to E$ is a random variable which is measurable with respect to $\mathcal{F}$ for each $t \in \mathcal{T}$. For each $\omega \in \Omega$, the process $X(\cdot, \omega) : \mathcal{T} \to E$ is called a realization, a trajectory or a sample path of the stochastic process. Often the dependency on Ω is suppressed and the stochastic process is written as $X(t)$ instead of $X(t, \cdot)$.

Let $\mathcal{F}_t$ be the σ-algebra which represents what is known up to time $t \in \mathcal{T}$. A collection of σ-algebras $\{\mathcal{F}_t\}$ such that $\mathcal{F}_t \subset \mathcal{F}$ for each t and $\mathcal{F}_s \subset \mathcal{F}_t$ if $s < t$ is called a filtration, which can be understood as a model for the evolution of available information over time. A filtration is right continuous if $\mathcal{F}_{t+} = \cap_{\epsilon>0}\mathcal{F}_{t+\epsilon} = \mathcal{F}_t$ for all $t > 0$ and it is complete if each $\mathcal{F}_t$ contains every null set. A filtration is said to satisfy the usual conditions if it is right continuous and complete.

A stochastic process is said to be adapted to a filtration $\{\mathcal{F}_t\}$ if it is $\mathcal{F}_t$ measurable for each t. That is, if X_t is adapted to $\{\mathcal{F}_t\}$, then the value of X_t is known at time t whenever the information represented by $\{\mathcal{F}_t\}$ is known.

The probabilistic behaviour of a stochastic process is driven by the assumptions on the random variables and the dependence between them. The dependence is defined by a probabilistic nature inherited on the random variables, for example in the form of probability distribution or described by the Markov property.

In the following, the definition of SDEs which generate stochastic processes is given.

6.1 SDEs

An SDE can be modelled with different kind of stochastic behaviour, which refers to the stochastic process which drives the SDE. Well-known processes influencing SDEs are the Wiener process, the Poisson point process and the Poisson random measure [2]. In the following, the presentation is limited to cover the case when the Wiener process drives the SDEs. The Wiener process is characterized by the Gaussian distribution and defined as follows.

Definition 4 (The standard Wiener process). *Let $W(t) = W_t$ be a one-dimensional $\mathbb{R}$-valued stochastic process with respect to the filtration $\{\mathcal{F}_t\}$ with the following properties:*

(1) *$W(t)$ is $\mathcal{F}_t$ measurable for each $t \geq 0$.*

(2) *$W(0) = 0$.*

(3) *The process has independent increment, that is, $W(t) - W(s)$ is independent of $\mathcal{F}_s$ whenever $s < t$.*

(4) *The process $W(t) - W(s)$ is Gaussian distributed with expected value zero and variance $t - s$ whenever $s < t$.*

(5) *$W(t)$ has continuous sample path.*

Then $W(t)$ is called a standard Wiener process [35].

The initial value of a standard Wiener process is defined to be zero. However, the process $\tilde{W}_t = W_t + c$, $c \in \mathbb{R}$, is also a Wiener process, which has properties (1),(3)-(5) from definition 4 but with $\tilde{W}_0 = c$.

The Wiener process models white noise behaviour and, therefore, it is particularly useful in modelling of real-world phenomena since observations of error contributions caused by many different factors are likely to be normal distributed. One of the main characteristic of the Wiener process is the fact that sample path of the process is everywhere continuous and nowhere differentiable with probability one. A proof of this important property can be found in [35].

The Wiener process is the stochastic term which drives the SDEs. Following, a multi-dimensional SDE is defined.

Definition 5 (A multi-dimensional SDE). *Let $\mathbf{X} \in \mathbb{R}^n$ be a stochastic process defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. A general multidimensional SDE is given by*

$$d\mathbf{X}_t = b(t, \mathbf{X}_t)dt + \sigma(t, \mathbf{X}_t)d\mathbf{W}_t, \quad \mathbf{X}_0 = \mathbf{c} \tag{9}$$

where

- $\mathbf{X}_t = \mathbf{X}(t) : \mathcal{T} \to \mathbb{R}^n$ *is an $\mathbb{R}^n$-valued stochastic process,*
- $b : \mathcal{T} \times \mathbb{R}^n \to \mathbb{R}^n$ *is the drift coefficient,*
- $\sigma : \mathcal{T} \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$ *is the diffusion coefficient,*
- $\mathbf{W}_t = (W_1(t), W_2(t), \ldots, W_d(t))^\top$ *is a d-dimensional Wiener process,*
- $\mathbf{c}$ *is a random variable independent of $\mathbf{W}_t - \mathbf{W}_0$ for $t \geq 0$ and with $\mathbb{E}[|\mathbf{c}|^2] < \infty$.*

The expression of the SDE given in (9) has to be understood as notation (a shorthand form of the integral equation), since the Wiener process is nowhere differentiable. The integral version of the SDE is

$$\mathbf{X}_t = \mathbf{c} + \int_0^t b(s, \mathbf{X}_s)ds + \int_0^t \sigma(s, \mathbf{X}_s)d\mathbf{W}_s \tag{10}$$

where the last term is the stochastic Itô integral, which is defined in section 6.2. A stochastic process generated by (10) is also called an Itô process [3].

Solutions to SDEs

The solution to an SDE is a stochastic process. The following theorem presents the general statement about existence and uniqueness of solutions to SDEs. The theorem is adapted from [3, Theorem 5.2.1].

Theorem 6 (Existence and uniqueness of solutions to SDEs). *For $t \in \mathcal{T}$ let the drift and diffusion be measurable functions satisfying the growth bound*

$$|b(t, \mathbf{X})| + |\sigma(t, \mathbf{X})| \leq C(1 + |\mathbf{X}|), \quad \mathbf{X} \in \mathbb{R}^n, t \in \mathcal{T}$$

and the Lipschitz condition

$$|b(t,\mathbf{X}) - b(t,\mathbf{Y})| + |\sigma(t,\mathbf{X}) - \sigma(t,\mathbf{Y})| \leq D|\mathbf{X}-\mathbf{Y}|, \quad \mathbf{X},\mathbf{Y} \in \mathbb{R}^n, t \in \mathcal{T}$$

for some constants C, D. Then the SDE in (9) *has a unique continuous solution $\mathbf{X}_t(\omega)$ for $t \in \mathcal{T}$ which is adapted to the filtration $\{\mathcal{F}_t\}$ generated by* $\mathbf{c}$ *and $\mathbf{W}_s$ for $s \leq t$. Furthermore,*

$$\mathbb{E}\left[\int_{\mathcal{T}} |\mathbf{X}_t|^2 dt\right] < \infty \,.$$

A solution to (9) is sometimes referred to as the pair of processes $(\mathbf{X}_t, \mathbf{W}_t)$. Depending on the information available to the process, the solution can be classified as weak or strong. The solution defined in Theorem 6 is a strong solution.

Definition 7 (Weak and Strong solutions). *Consider the SDE presented in definition 5 and let $\{\mathcal{F}_t\}$ be the filtration generated by $\mathcal{F}$ which satisfies the usual conditions.*

- *A weak solution $(\mathbf{X}, \mathbf{W})$ to* (9) *exists if $\mathbf{X}_t$ is a continuous $\{\mathcal{F}_t\}$ adapted process with values in $\mathbb{R}^n$ such that* (9) *holds [35].*
- *Furthermore, a solution $(\mathbf{X}, \mathbf{W})$ is said to be a strong solution to* (9) *if for all $t > 0$ the process $\mathbf{X}_t$ is adapted to the filtration $\{\mathcal{F}_t^W\}$ generated by $\mathbf{W}_s$ for $s \leq t$ [39].*

When the solution exists, one of the core challenges in modelling with SDEs is to find an explicit expression of the solution which solves the SDE. An explicit solution is defined in the following way.

Definition 8 (An Explicit Solution). *A solution of an SDE is called an explicit solution, if it has a representation where the terms on its right hand side do not use the solution itself [2].*

Given an explicit solution, it is relatively easy to obtain realizations or sample paths of the stochastic process by generating the process from its initial condition. In other cases, realizations can be simulated by applying numerical methods, which is discussed in section 7.

The stochastic integral included in (10) has to be understood as the Itô integral, which is presented in the following together with a general presentation of Itô calculus.

6.2 Itô Calculus

Itô calculus is a method which extends ordinary calculus methods to calculus with stochastic processes. Following, the Itô integral is defined.

Definition 9 (The Itô integral). *Suppose that W is a one-dimensional standard Wiener process and that $\sigma(t,\omega)$ is*

(p1) measurable in t,

(p2) $\mathcal{F}_t$*-adapted*

(p3) and that $\mathbb{E}[\int_0^t \sigma(t,\omega)^2 dt] < \infty$.

Let the time interval $\mathcal{T} = [0,T]$ *be partitioned into n sub-intervals by the time instants* $0 = t_0, t_1, \ldots, t_n = T$. *Then the Itô integral is defined by*

$$\int_0^T \sigma(t,\omega) dW_t \quad = \quad \lim_{n\to\infty} \sum_{i=1}^{n} \sigma(t_{i-1},\omega)(W_{t_i} - W_{t_{i-1}}) \; . \tag{11}$$

The Itô integral has a number of interesting properties. If $\sigma(t,\omega)$ has properties $(p1) - (p3)$ from definition 9, then the expected value of the Itô integral is zero, i.e.

$$\mathbb{E}\left[\int_0^T \sigma(t,\omega) dW_t\right] \quad = \quad 0 \; . \tag{12}$$

This result follows directly by the definition of the Itô integral and property (4) from definition 4.

Another useful property is the Itô isometry which says the following.

Proposition 10 (The Itô isometry). *If* $\sigma(t,\omega)$ *has properties* $(p1) - (p3)$ *definition 9, then*

$$\mathbb{E}\left[\left(\int_s^T \sigma(t,\omega) dW_t(\omega)\right)^2\right] \quad = \quad \mathbb{E}\left[\int_s^T \sigma(t,\omega)^2 dt\right] \; .$$

[3, Corollary 3.1.7]

In cases where an explicit expression of the solution to an SDE exists, it can be isolated by applying Itô's formula, which corresponds the the chain rule for ordinary calculus.

Theorem 11 (Itô's formula). *Let* $\mathbf{X}(t) = (X_1, \ldots, X_n)$ *be a stochastic process given by the SDE in* (9). *Let* $g(t,x) = (g_1(t,x), \ldots, g_p(t,x))$ *be a* C^2 *map from* $[0,\infty) \times \mathbb{R}^n$ *to* $\mathbb{R}^p$. *Then the process*

$$Y(t,\mathbf{X}) \quad = \quad g(t,\mathbf{X}(t))$$

is an Itô process, whose component number k, Y_k *is given by*

$$dY_k \quad = \quad \frac{\partial g_k}{\partial t}(t,\mathbf{X})dt + \sum_i \frac{\partial g_k}{\partial x_i}(t,\mathbf{X})dX_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,\mathbf{X})dX_i dX_j$$

where the multiplications are computed according to the rules

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0, \quad \textit{and} \quad dW_s \cdot dW_t = \begin{cases} dt & \textit{if } s = t \\ 0 & \textit{if } s \neq t \end{cases}$$

where W is a one-dimensional Wiener process [3, Theorem 4.2.1].

The following example is an SDE which solution can be expressed explicitly by applying Itô's formula.

Example 6.1 (The one-dimensional Ornstein-Uhlenbeck (OU) Process)
Consider the SDE given by

$$dX_t = \theta\left(\mu - X_t\right)dt + \sigma dW_t, \quad X(0) = x_0 \ .$$

The solution to this SDE is called the Ornstein-Uhlenbeck (OU) Process after G. E. Uhlenbeck and L. S. Ornstein who introduced the SDE as a model for the velocity of Brownian particles [50].

Following, let $\theta = \frac{1}{2}$, $\mu = 0$ and $\sigma = 1$ such that the SDE becomes

$$dX_t = -\frac{1}{2}X_t dt + dW_t, \quad X(0) = x_0 \ . \tag{13}$$

Let $Y(t,x) = xe^{\frac{t}{2}}$. Then Itô's formula gives

$$dY \quad = \quad \frac{1}{2}Xe^{\frac{t}{2}}dt + e^{\frac{t}{2}}\left(-\frac{1}{2}Xdt + dW_t\right) \quad = \quad e^{\frac{t}{2}}dW_t$$

and it follows that

$$Xe^{\frac{t}{2}} \quad = \quad x_0 + \int_0^t e^{\frac{s}{2}}dW_s$$

which means that the explicit solution to (13) is

$$X_t = x_0 e^{-\frac{t}{2}} + e^{-\frac{t}{2}}\int_0^t e^{\frac{s}{2}}dW_s \ . \tag{14}$$

The Itô isometry and the fact that the expectation of the Itô integral is zero can be applied to calculate the expectation, the variance and the covariance of the OU process.

Example 6.2 (The OU Process)
Consider the OU process X_t which is given in (14). The expected value of the OU process is

$$\mathbb{E}[X_t] \quad = \quad \mathbb{E}[x_0 e^{-\frac{t}{2}} + e^{-\frac{t}{2}}\int_0^t e^{\frac{s}{2}}dW_s] \quad = \quad x_0 e^{-\frac{t}{2}}$$

due to (12). Furthermore, the variance of the OU-process is

$$\begin{aligned}
\mathrm{Var}[X_t] &= \mathbb{E}[(x_0 e^{-\frac{t}{2}} + e^{-\frac{t}{2}} \int_0^t e^{\frac{s}{2}} dW_s - x_0 e^{-\frac{t}{2}})^2] \\
&= \mathbb{E}[(e^{-\frac{t}{2}} \int_0^t e^{\frac{s}{2}} dW_s)^2] \quad = \quad e^{-t}\mathbb{E}[(\int_0^t e^{\frac{s}{2}} dW_s)^2] \\
&= e^{-t}\mathbb{E}[\int_0^t (e^{\frac{s}{2}})^2 ds] \quad = \quad e^{-t}\mathbb{E}[\int_0^t e^s ds] \\
&= e^{-t} \int_0^t e^s ds \quad = \quad 1 - e^{-t}
\end{aligned}$$

where the Itô isometry is applied. The covariance of X_t and X_s is

$$\begin{aligned}
Cov[X_t, X_s] &= \mathbb{E}[(X_t - x_0 e^{-\frac{t}{2}})(X_s - x_0 e^{-\frac{s}{2}})] \\
&= \mathbb{E}[(e^{-\frac{t}{2}} \int_0^t e^{\frac{r}{2}} dW_r)(e^{-\frac{s}{2}} \int_0^s e^{\frac{r}{2}} dW_r)] \\
&= e^{-\frac{t}{2}} e^{-\frac{s}{2}} \mathbb{E}[\int_0^t e^{\frac{r}{2}} dW_r \int_0^s e^{\frac{r}{2}} dW_r] \quad = \quad e^{-\frac{s+t}{2}} \mathbb{E}[\int_0^{s \wedge t} e^r dr] \\
&= e^{-\frac{s+t}{2}} \left(e^{s \wedge t} - 1\right)
\end{aligned}$$

where $s \wedge t = \min(s, t)$.

Relation to the Stratonovich integral

The Itô integral is one definition of the stochastic integral. Another definition is made by Stratonovich and there is a direct relation between the two definitions.

The Stratonovich integral is defined as the limit in mean square by

$$\int_0^t \sigma(s, \omega) \circ dW_s \quad = \quad \lim_{n \to \infty} \sum_{i=0}^{n-1} \frac{\sigma(t_{i+1}, \omega) + \sigma(t_i, \omega)}{2} (W_{t_{i+1}} - W_{t_i}) \, .$$

The transformation between the Stratonovich integral and the Itô integral defined in (11) is

$$\int_0^t \sigma(s, X_s) \circ dW_s \quad = \quad \frac{1}{2} \int_0^t \sigma'(s, X_s)\sigma(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s$$

where $'$ denotes the derivative of $\sigma(s, X)$ with respect to X [3]. Notice that this transformation only holds true for SDEs with differentiable diffusion coefficient.

The stochastic evolution in time can be described by the Kolmogorov forward and backward equations, which characterize the evolution of the stochastic process in probability. That is, the equations describe the probability that the stochastic process is in a certain state at a certain time and how it changes over time [40]. The forward

Kolmogorov equation is also known as the Fokker-Planck equation, which is presented in the following.

6.3 The Fokker-Planck Equation

The dynamics of a system can be described probabilistically with the Fokker-Planck equation. This method was used by Fokker and Planck to describe the Brownian motion of particles and, therefore, the Fokker-Planck equation is also suitable to describe systems driven by SDEs [51].

The Fokker-Planck equation describes the probabilistic evolution of a stochastic process over time. That is, the solution to a Fokker-Planck equation has the properties of a (probability) density function.

Definition 12 (The Fokker-Planck equation). *Let $p(\mathbf{X}, t)$ be the (probability) density function for a stochastic process driven by an SDE defined as in* (9)*. The Fokker-Planck equation gives the following relation*

$$\frac{\partial}{\partial t}p(t,\mathbf{X}) = -\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[b_i(t,\mathbf{X})p(t,\mathbf{X})\right] + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2}{\partial x_i \partial x_j}\left[D_{i,j}(t,\mathbf{X})p(t,\mathbf{X})\right]$$

where

$$D_{i,j}(t,\mathbf{X}) \quad = \quad \sum_{k=1}^{d}\sigma_{ik}(t,\mathbf{X})\sigma_{jk}(t,\mathbf{X})\ . \tag{15}$$

In general, it is very difficult to obtain a solution to the Fokker-Planck equation so there are developed many numerical methods to generate approximated solutions. In some cases, the exact solution can be found by transformation methods. An example of this is presented in section 7.2 where the method of Fourier transformation is applied to determine the solution to a Fokker-Planck equation.

The solutions to the Fokker-Planck equation can be time-dependent, but in many cases a time-invariant solution also exists, which is called a stationary solution. Let $p(\mathbf{X})$ be the stationary solution to the Fokker-Planck equation. Then the Fokker-Planck equation becomes

$$0 \quad = \quad -\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left[b_i(t,\mathbf{X})p(\mathbf{X})\right] + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2}{\partial x_i \partial x_j}\left[D_{i,j}(t,\mathbf{X})p(\mathbf{X})\right]$$

where $D_{i,j}(t,\mathbf{X})$ is given in (15).

In the following example the stationary solution to the Fokker-Planck equation for the OU process is determined.

Example 6.3 (The stationary Fokker-Planck equation for the OU process)
The stationary Fokker-Planck equation for OU process defined in (13) is

$$0 = \frac{1}{2}\frac{\partial}{\partial x}xf(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(x) = \frac{1}{2}f(x) + \frac{1}{2}x\frac{\partial}{\partial x}f(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(x)\,. \qquad (16)$$

A solution to this equation takes the form ce^{-ax^2}, where a and c are constants. By substituting this expression into (16) and differentiating accordingly, the stationary solution is obtained to be

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$$

where the constant c is determined under the assumption that $\int_{\mathbb{R}} f(x)dx = 1$.

7 Numerical Methods

The solution to SDEs can be hard or impossible to express explicitly. Instead, realizations of solutions can be estimated by application of numerical methods. The same counts for the Fokker-Planck equation. With numerical methods it is also possible to approximate the probability solution to this equation. In the following, first a numerical method which applies to SDEs is presented and afterwards numerical- and transformation techniques applicable to the Fokker-Planck equation are discussed.

7.1 The Euler-Maruyama Method

The Euler-Maruyama method is a stochastic extension of the well-known Euler method for differential equations [42]. The drift and diffusion coefficient from a given SDE are implemented into a stochastic recursion to generate numerical realizations of the stochastic process driven by the original SDE. In the following, the presentation is limited to the one-dimensional case, but the method is also valid for multidimensional SDEs.

Consider the autonomous one-dimensional SDE given in integral form

$$X(t) \quad = \quad X(t_0) + \int_{t_0}^{t} b(X(s))ds + \int_{t_0}^{t} \sigma(X(s))dW_s$$

where $b, \sigma : \mathbb{R} \to \mathbb{R}$ are drift and diffusion coefficient and W_t is the one-dimensional standard Wiener process.

A numerical solution $\{Y_n\}$ is generated over fixed time interval $\mathcal{T} = [t_0, T]$ with the initial position $Y_0 = X(t_0)$. Thus, the time interval is partitioned into N subintervals by equispaced points such that $t_0 < t_1 < \ldots < t_N$ with the length of each time interval given by $\Delta t = t_{i+1} - t_i = \frac{T-t_0}{N}$ for $i \in \{0, \ldots, N-1\}$.

For each recursive step, that is, for $t \in [t_i, t_{i+1}]$ for $i \in \{0, \ldots, N-1\}$, the influence of the Wiener process has to be observed. Therefore, the stochastic increment of the Wiener process given by

$$\Delta W_i \quad = \quad W_{t_{i+1}} - W_{t_i}$$

for $i \in \{0, \ldots, N-1\}$ is included. Recall that the Wiener process has the property that the increments ΔW_i are independent and identically distributed normal variables with expected value zero and variance Δt (see definition 4). Therefore, for each recursive step ΔW_i can be generated from the Gaussian distribution.

Then the approximated solution $\{Y_n\}$ in the time interval $\mathcal{T} = [t_0, T]$ can be generated for $n \in \{0, \ldots, N-1\}$ by the recursion

$$Y_{n+1} \quad = \quad Y_n + b(Y_n)\Delta t + \sigma(Y_n)\Delta W_n, \quad Y_0 = X(t_0)$$

where $b(\cdot)$ and $\sigma(\cdot)$ are drift and diffusion coefficient of the original SDE, respectively.

Example 7.1 (The OU process)
A realization $\{Y_n\}$ of the OU process given in (13) is simulated with the Euler-Maruyama method with $\Delta t = 2^{-9}$. The result is compared with the explicit solution given by (14) in Figure 12.

Rate of convergence

The quality of a numerical method can be validated by considering the rate of convergence, which is categorised in weak and strong convergence.

Definition 13 (Strong and weak rate of convergence). *Let X be a stochastic process and let Y be an approximated solution to X generated with an approximation method where the discrete time step has length Δt. The approximation method is said to converge strongly with order $\gamma > 0$ at time $t \in \mathcal{T}$ if there exists a positive constant C which does not depend on Δt such that*

$$\mathbb{E}[|X(t) - Y(t)|] \quad \leq \quad C\Delta t^{\gamma} .$$

Furthermore, the discrete approximation Y is said to converge weakly with order $\beta > 0$ to X at time $t \in \mathcal{T}$ if there exists a positive constant C which does not depend on Δt such that

$$|\mathbb{E}[X(t)] - \mathbb{E}[Y(t)]| \quad \leq \quad C\Delta t^{\beta} .$$

[40]

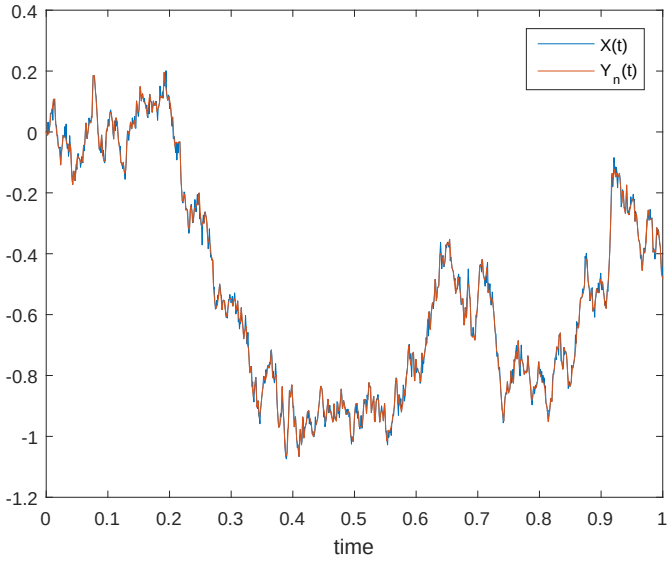

Figure 12: The explicit solution $X(t)$ to the OU process together with a realization $Y_n(t)$ simulated with the Euler-Maruyama method for $t \in [0, 1]$.

The rate of convergence is highly dependent on the drift and diffusion coefficient of the SDE which solution is approximated. The Euler-Maruyama method converges strongly with order $\gamma = \frac{1}{2}$ and weakly with order $\beta = 1$ when both the drift and diffusion coefficient are constants. These are the best strong and weak orders which can be obtained for the Euler-Maruyama method.

In paper B the Euler-Maruyama method is applied to an SDE with discontinuous drift coefficient and it is proved that the constructed realizations converge to the strong solution of the original SDE. In this case, the strong rate of convergence is obtained to be $\gamma = \frac{1}{4}$.

7.2 Solution Methods to the Fokker-Planck Equation

The time-dependent solution to the Fokker-Planck equation can be hard to solve analytically, but alternative methods are developed. Following, the discrete Fourier series is presented and the application to the Fokker-Planck equation is outlined.

The discrete Fourier series

Fourier analysis is the process of representing or approximating functions with sinusoidal waveforms [52]. This method can often simplify the study of the original function. In this section, the discrete Fourier transform is applied when it is developed

over the basis

$$\left\{1, \cos\left(\frac{\pi n x}{L}\right), \sin\left(\frac{\pi n x}{L}\right) : n \in \{1, 2, 3, \ldots\}\right\}$$

for $x \in [-L, L]$, $L \in \mathbb{N}$. A real-valued function $f(x)$ can be expressed by the given basis for $x \in [-L, L]$ by

$$f(x) = \frac{a}{2} + \sum_{n=1}^{\infty} b_n \cos\left(\frac{\pi n x}{L}\right) + \sum_{n=1}^{\infty} c_n \sin\left(\frac{\pi n x}{L}\right) \tag{17}$$

where

$$\begin{aligned} a &= \frac{1}{L}\int_{-L}^{L} f(x)dx \,, \\ b_n &= \frac{1}{L}\int_{-L}^{L} f(x)\cos\left(\frac{\pi n x}{L}\right)dx \,, \\ c_n &= \frac{1}{L}\int_{-L}^{L} f(x)\sin\left(\frac{\pi n x}{L}\right)dx \,. \end{aligned}$$

In practice, the expression in (17) requires considerable effort to evaluate and, therefore, only an approximation of $f(x)$ can be determined. Thus, for sufficiently big M the discrete Fourier series can be applied to approximate a function defined over a limited interval by truncating the sum expressions. That is,

$$f(x) \approx \frac{a}{2} + \sum_{n=1}^{M} b_n \cos\left(\frac{\pi n x}{L}\right) + \sum_{n=1}^{M} c_n \sin\left(\frac{\pi n x}{L}\right) \,,$$

for sufficiently big M. This approximation method can be used to obtain an approximated solution to the Fokker-Planck equation. Therefore, consider the Fokker-Planck equation for a one-dimensional SDE,

$$\frac{\partial}{\partial t}p(t,x) = -\frac{\partial}{\partial x}\left[b(t,x)p(t,x)\right] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[\sigma^2(t,x)p(t,x)\right] \,.$$

Let $\tilde{p}(t,x)$ be an approximation for $x \in [-L, L]$ of the solution $p(t,x)$ to the Fokker-Planck equation given by

$$\tilde{p}(t,x) = \frac{a(t)}{2} + \sum_{n=1}^{M} b_n(t)\cos\left(\frac{\pi n x}{L}\right) + \sum_{n=1}^{M} c_n(t)\sin\left(\frac{\pi n x}{L}\right) \,,$$

for sufficiently big M and with

$$\begin{aligned} a(t) &= \frac{1}{L}\int_{-L}^{L} \tilde{p}(t,x)dx \,, \\ b_n(t) &= \frac{1}{L}\int_{-L}^{L} \tilde{p}(t,x)\cos\left(\frac{\pi n x}{L}\right)dx \,, \\ c_n(t) &= \frac{1}{L}\int_{-L}^{L} \tilde{p}(t,x)\sin\left(\frac{\pi n x}{L}\right)dx \,. \end{aligned}$$

Substitution of $\tilde{p}(t,x)$ into the Fokker-Planck equation gives the expression

$$\frac{\partial}{\partial t}\frac{a(t)}{2} + \sum_{n=1}^{M}\frac{\partial}{\partial t}b_n(t)\cos\left(\frac{\pi n x}{L}\right) + \sum_{n=1}^{M}\frac{\partial}{\partial t}c_n(t)\sin\left(\frac{\pi n x}{L}\right) \tag{18a}$$

$$= -\frac{\partial}{\partial x}\Big[b(t,x)\frac{a(t)}{2} + b(t,x)\sum_{n=1}^{M}b_n(t)\cos\left(\frac{\pi n x}{L}\right) \tag{18b}$$

$$+ b(t,x)\sum_{n=1}^{M}c_n(t)\sin\left(\frac{\pi n x}{L}\right)\Big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big[\sigma^2(t,x)\frac{a(t)}{2} \tag{18c}$$

$$+ \sigma^2(t,x)\sum_{n=1}^{M}b_n(t)\cos\left(\frac{\pi n x}{L}\right) + \sigma^2(t,x)\sum_{n=1}^{M}c_n(t)\sin\left(\frac{\pi n x}{L}\right)\Big]. \tag{18d}$$

To obtain the approximation $\tilde{p}(t,x)$, it is necessary to determine the coefficients

$$a(t), b_1(t), \ldots, b_M(t), c_1(t), \ldots, c_M(t)\ ,$$

which can be done by solving the system of ordinary differential equations appearing in (18).

This method of using discrete Fourier series to estimate functions is applied in paper D to approximate conditional density functions.

Laplace and Fourier transform

The Laplace and Fourier transform are both integral transformation methods which take real-valued functions to complex valued functions. The Laplace transform $L(f)(s)$ of a function f is a complex function of a complex variable s, while the Fourier transform $\hat{f}(k)$ of a function f is complex function of a real variable k. Following, the methods are defined.

Definition 14 (Laplace and Fourier transform). *The Laplace transform of an arbitrary function $f(x)$ is defined as*

$$L(f)(s) \quad = \quad \int_0^\infty f(x)e^{-sx}dx$$

where s is a complex number.

The Fourier transform of an arbitrary function $f(x)$ is defined as

$$\hat{f}(k) \quad = \quad \int_{-\infty}^{\infty} e^{-ikx}f(x)dx$$

where $i^2 = -1$.

The methods of Laplace and Fourier transform are applied in paper A to investigate the evolution of recursively defined density functions.

The Fourier transformation can also be applied to determine the time-dependent solution to the Fokker-Planck equation exactly. In this process, the inverse Fourier transform is needed, which is defined by

$$f(x) \quad = \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\hat{f}(k)dk \; .$$

As an example, the time-dependent solution to the Fokker-Planck equation for the OU process is presented.

Example 7.2 (The OU process)
The Fokker-Planck equation for the OU process presented in example 6.1 is

$$\frac{\partial}{\partial t}f(t,x) \quad = \quad \frac{1}{2}\frac{\partial}{\partial x}xf(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(t,x) \; .$$

Let $\hat{f}(k,t)$ be the Fourier transform of $f(t,x)$. Then the Fourier transform of the terms in the Fokker-Planck equation becomes

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{ikx}\frac{\partial}{\partial t}f(t,x)dx &= \frac{\partial}{\partial t}\hat{f}(k,t) \\
\int_{-\infty}^{\infty} e^{ikx}\frac{1}{2}\frac{\partial}{\partial x}xf(t,x)dx &= -\frac{1}{2}k\int_{-\infty}^{\infty} e^{ikx}ixf(t,x)dx = -\frac{1}{2}k\frac{\partial}{\partial k}\hat{f}(k,t) \\
\int_{-\infty}^{\infty} e^{ikx}\frac{1}{2}\frac{\partial^2}{\partial x^2}f(t,x)dx &= -\frac{1}{2}k^2\hat{f}(k,t) \; .
\end{aligned}$$

Therefore, the Fokker-Planck equation gives the relation in the Fourier domain

$$\frac{\partial}{\partial t}\hat{f}(k,t) \quad = \quad -\frac{1}{2}k\frac{\partial}{\partial k}\hat{f}(k,t) - \frac{1}{2}k^2\hat{f}(k,t) \; . \tag{19}$$

Let the initial condition at time $t = 0$ be a constant x_0 such that $f(x,0) = \delta(x - x_0)$. The Fourier transform of the initial condition becomes

$$\hat{f}(k,0) \quad = \quad e^{-ikx_0} \; .$$

Then the solution to (19) can be determined by using the first order method of characteristics for PDEs [53]. This gives the result

$$\hat{f}(k,t) \quad = \quad \exp\left(-ikx_0e^{-\frac{1}{2}t} - \frac{1}{2}k^2\left(1 - e^{-t}\right)\right) \; .$$

Applying the inverse Fourier transformation gives the solution to the original Fokker-Planck equation,

$$f(t,x) \quad = \quad \frac{1}{\sqrt{2\pi\left(1 - e^{-t}\right)}}\exp\left(-\frac{(x - x_0e^{-\frac{1}{2}t})^2}{2(1 - e^{-t})}\right) \; .$$

Note that as $t \to \infty$, the function approaches the stationary solution presented in example 6.3.

Summary of Contributions

This chapter summarizes the main contributions of the PhD project.

8 A Simple SDE with Discontinuous Drift

This PhD project was initiated by studying of a simple SDE with discontinuous drift. This was motivated by a step-by-step approach. By understanding the behaviour of the solution to a simple one-dimensional SDE with discontinuous drift, it might be possible to extend this insight to other classes of SDEs in later approaches.

The results in paper A are founded by a numerical and heuristics approach to study the solutions to a simple SDE with discontinuous drift. Two approaches are applied: The Euler-Maruyama method and the Fokker-Planck equation.

The Euler-Maruyama method is basis for a density function approach, which results in recursively defined density functions for the position of the stochastic process at certain time instants. The density functions and their probabilistic properties are shown to approach stationary behaviour as the step-length of the Euler-Maruyama method decreases.

The Fokker-Planck equation is applied to produce a stationary density function for the solution to the simple SDE with discontinuous drift. A key observation in the paper is that the density functions produced by the Euler-Maruyama method approximate the stationary density function generated by the Fokker-Planck equation.

In order to compensate for the discontinuous challenges and/or to compare with standard results, paper A also introduces a smooth function which approximates the discontinuous function in the area of the discontinuous point. The Euler-Maruyama method and the Fokker-Planck procedure are repeated with this smooth function. The numerical results do not seem to distinguish significantly from the results obtained with the discontinuous function, and the stationary solution to the Fokker-Planck equation does clearly approach the result obtained with discontinuous dynamics as the smooth function approaches the discontinuous function.

The correctness of applying the Euler-Maruyama method to SDEs with discontinuous drift is validated in paper B.

Paper B investigates convergence properties of the Euler-Maruyama method when it is applied to the same simple SDE with discontinuous drift as the one investigated

in paper A. It is proven that numerical solutions produced with the Euler-Maruyama method converge to the strong solution of the SDE (which exists due to [39]). Furthermore, the strong rate of convergence of the Euler-Maruyama method is obtained to be at least equal to $\frac{1}{4}$. This result indicates that the Euler-Maruyama method has the potential to be applied as a method to produce numerical solutions to general versions of SDEs with discontinuous drift coefficient.

The discontinuous drift introduced in paper A and B has its characteristic from the sgn-function. Many practical applications of sliding mode control apply the same discontinuous function in the control design. Therefore, it seems natural to continue the direction of the research by investigating the switching dynamics in sliding mode control.

9 The Switching Dynamics in Sliding Mode Control

In sliding mode control the designed controller produces a switching mechanism based on discontinuous dynamics. This switching mechanism is investigated in paper C while a complete description of the system with a sliding mode controller is analysed in paper D.

Paper C investigates the switching dynamics in sliding mode control under the assumption that the switching is driven by a Poisson point process. The analysis of the original sliding mode approach is simplified by a transformation of the state space, such that only a small region of the state space is considered. A probabilistic model is developed which results in an explicit description of the marginal density functions for the stochastic process which is generated by the switching mechanism.

In paper D a sliding mode controller is applied to a mechanical system which is exposed to additive white noise. A linear coordinate transformation of the system is introduced which implies that known results on SDEs with discontinuous drift can be applied to the system. The transformed system is analysed with Itô calculus which includes a mean valued analysis of the stochastic Itô integral.

The main contribution of paper D is a statistical characterization of system performance in terms of the stationary variance of the control error. In order to determine this result, the auto-correlation function for one part of the stochastic process is approximated with discrete Fourier series.

The practical relevance of the results in both papers C and D is situated in the potential to describe or estimate the collected behaviour of a family of physical systems connected via certain switching laws governed by some degree of randomness or by some version of sliding mode control. The results in paper C can be used for local analysis of such systems and the results in paper D provide information on the system behaviour near the system's operating point.

Conclusion and Perspectives

This chapter concludes the research project and introduces possible directions for future investigations

10 Models of Stochastic Systems with Switching Dynamics

During the studies of this PhD project, stochastic switching dynamics was investigated from different perspectives by different models and methods. The underlying goals was the same all the way; to increase knowledge and insight in how to interpret and understand the behaviour generated by models of real-world phenomena/systems which are exposed to discontinuous dynamics.

The results in paper B indicate that the Euler-Maruyama method has the potential to be applied as a method to produce numerical solutions to SDEs with discontinuous drift coefficient. However, the method needs also to be approved for more general classes of SDEs with discontinuous drift coefficient. This is, partly, done in the novel work [49], which proves strong convergence of the Euler-Maruyama method for the case where both the drift and diffusion coefficient are discontinuous. However, also in this work, the drift and diffusion coefficient are restricted to be bounded. Therefore, it is still of interest to investigate whether the Euler-Maruyama method is suitable for constructing numerical solutions to SDEs with unbounded discontinuous drift coefficient.

Paper D makes the initial attempt to study SDEs with unbounded discontinuous drift and provides reliable estimates on the system behaviour for the class of (mechanical) systems where sliding mode control is applied. The perspective of this work is to extend the method to more general classes such that other systems can also be modelled with SDEs even though they are subject to discontinuous dynamics.

The analysis of SDEs with unbounded discontinuous drift is a key issue that is still poorly understood after completion of this thesis.

Most of the results obtained in paper A, C and D have stationary characteristics. A next step is to investigate the time-dependent evolution of the system's performance. As an initial attempt, the following model and procedure is suggested.

A stochastic model with Markovian switching

One of the methods investigated in paper A applies the Fourier transform to recursively generated density functions. The idea is to introduce the transformation before the convolution of density functions such that the convolution procedure can be substituted with multiplication. Even though this method did not bring any significant result in paper A, this transformation method might be useful for other models. Hence, a new model is formulated below and the method of Fourier transformation is applied.

First, recall the model presented in paper C. This model defines two constant vector fields, which drives the trajectory towards an equilibrium point. The switching dynamics between the two vector fields is modelled with the Poisson point process. The switching mechanism is active when the trajectory has passed the switching surface which implies that the switching instant is always subject to some (state space) delay.

Now, let the model in paper C be extended to the framework of SDEs and the dynamics of the system be considered separated, but interconnected, that is, as a kind of stochastic hybrid model. Let $v_1 > 0$ and $v_2 > 0$ denote the constant vector fields and β be the diffusion coefficient of the SDEs. Then, the dynamics of this system can be expressed by the two dependent Fokker-Planck equations given by

$$\begin{aligned}
\frac{\partial \phi_1}{\partial t} &= -v_1 \frac{\partial \phi_1}{\partial x} + \frac{\beta^2}{2} \frac{\partial^2 \phi_1}{\partial x^2} + \lambda \mathbb{I}_{(x<0)} \phi_2 - \mu \mathbb{I}_{(x>0)} \phi_1 \\
\frac{\partial \phi_2}{\partial t} &= v_2 \frac{\partial \phi_2}{\partial x} + \frac{\beta^2}{2} \frac{\partial^2 \phi_2}{\partial x^2} + \mu \mathbb{I}_{(x>0)} \phi_1 - \lambda \mathbb{I}_{(x<0)} \phi_2
\end{aligned}$$

where $\phi_1(t,x)$ and $\phi_2(t,x)$ are (probability) density functions for $x(t)$ in each of the two discrete states, and λ and μ correspond to the intensity of the Poisson point process (here modelled as Markovian switching).

One way to interpret the Fokker-Planck equations is to think of the stochastic process as $(x(t), q(t))$ where $x(t)$ is the continuous state and $q(t) \in \{1,2\}$ is the discrete state which defines the current active vector field. When the stochastic process is in the state $(x(t) < 0, q(t) = 1)$, the drift term drives the process in the direction towards $x = 0$ so the probabilistic evolution of $x(t)$, in this case given by $\phi_1(t,x)$, is not exposed to any (spontaneous) Markov switches (corresponding to the discrete transition generated by the Poisson point process in paper C). However, when $(x(t) < 0, q(t) = 2)$ the drift drives the state $x(t)$ away from $x = 0$ and the probability of a Markov switch is active. Therefore, in the evolution of the density functions, $\phi_2(t,x)$ risks to "lose" probability mass to $\phi_1(t,x)$ which gains probability mass. The opposite interpretation appears for $(x(t) > 0, q(t))$.

The question is whether the application of the Fourier (or Laplace) transform to the above Fokker-Planck equations will make it possible to obtain expressions for the time-dependent density functions $\phi_1(t,x)$ and $\phi_2(t,x)$.

So far, it is observed that the expectation of the stochastic process $x(t)$ is given by

$$E[x(t)] \quad = \quad E[x(0)] + \int_0^t v_1\hat{\phi}_1(\tau,0) - v_2\hat{\phi}_2(\tau,0)d\tau$$

where $\hat{\phi}_1$ and $\hat{\phi}_2$ are the Fourier transformed density functions. More thoughts and details on this method are presented in the Appendix, section 1.

References

[1] H. K. Khalil, *Nonlinear Systems*, Third, Ed. Prentice-Hall. Inc., 2000.

[2] E. Platen and N. Bruti-Liberati, *Numerical solution of stochastic differential equations with jumps in finance*, ser. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2010, vol. 64. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-13694-8

[3] B. Øksendal, *Stochastic differential equations*, 6th ed., ser. Universitext. Berlin: Springer-Verlag, 2003, an introduction with applications. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-14394-6

[4] J.-J. E. Slotine and W. Li, *Applied Nonlinear Control*. Englewood Cliffs, New Jersey 07632: Prentice Hall, 1991.

[5] T. Glad and L. Ljung, *Control Theory*. 11 New Fetter Lane, London EC4P 4EE: Taylor & Francis, 2000, multivariable and Nonlinear Methods.

[6] H. L. Trentelman, A. A. Stoorvogel, and M. Hautus, *Control theory for linear systems*, ser. Communications and Control Engineering Series. Springer-Verlag London, Ltd., London, 2001. [Online]. Available: http://dx.doi.org/10.1007/978-1-4471-0339-4

[7] D. Liberzon, *Switching in systems and control*, ser. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2003. [Online]. Available: http://dx.doi.org/10.1007/978-1-4612-0017-8

[8] Z. Sun and S. S. Ge, *Switched Linear Systems Control and Design*, ser. Communications and Control Engineering Series. Springer, 2005.

[9] J. a. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," *Decision and Control, 1999. CDC. 38th IEEE Conference on*, pp. 2655–2660, 1999.

[10] L. Vu, D. Chatterjee, and D. Liberzon, "Input-to-state stability of switched systems and switching adaptive control," *Automatica J. IFAC*, vol. 43, no. 4, pp. 639–646, 2007. [Online]. Available: http://dx.doi.org/10.1016/j.automatica.2006.10.007

[11] H. S. Witsenhausen, "A class of hybrid-state continuous-time dynamic systems," *IEEE Trans. on Auto. Cont.*, vol. Ac-11, no. 2, pp. 161–167, 1966.

[12] L. M. Bujorianu, *Stochastic Reachability Analysis of Hybrid Systems*. Springer-Verlag London Limited, 2012.

[13] R. Goebel, R. G. Sanfelice, and A. R. Teel, "Hybrid dynamical systems: robust stability and control for systems that combine continuous-time and discrete-time dynamics," *IEEE Control Syst. Mag.*, vol. 29, no. 2, pp. 28–93, 2009. [Online]. Available: http://dx.doi.org/10.1109/MCS.2008.931718

[14] ——, *Hybrid dynamical systems*. Princeton University Press, Princeton, NJ, 2012, modeling, stability, and robustness.

[15] R. Malhame and C.-Y. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Transactions on Automatic Control*, vol. AC-30, no. 9, pp. 852–860, 1985. [Online]. Available: http://dx.doi.org/10.1109/TAC.1985.1104071

[16] E. Altman and V. Gaitsgory, "Asymptotic optimization of a nonlinear hybrid system governed by a Markov decision process," *SIAM J. Control Optim.*, vol. 35, no. 6, pp. 2070–2085, 1997. [Online]. Available: http://dx.doi.org/10.1137/S0363012995279985

[17] J. A. Filar, V. Gaitsgory, and A. B. Haurie, "Control of singularly perturbed hybrid stochastic systems," *IEEE Trans. Automat. Control*, vol. 46, no. 2, pp. 179–190, 2001. [Online]. Available: http://dx.doi.org/10.1109/9.905686

[18] J. Hu, J. Lygeros, and S. Sastry, "Towards a theory of stochastic hybrid systems," *N. Lynch and B. Krogh (Eds.): HSCC 2000, LNCS 1790*, pp. 160–173, 2000.

[19] M. L. Bujorianu and J. Lygeros, "Toward a general theory of stochastic hybrid systems," in *Stochastic hybrid systems*, ser. Lecture Notes in Control and Inform. Sci. Berlin: Springer, 2006, vol. 337, pp. 3–30.

[20] J. Leth, J. G. Rasmussen, H. Schioler, and R. Wisniewski, "A class of stochastic hybrid systems with state-dependent switching noise," *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pp. 4737–4744, 2012. [Online]. Available: http://dx.doi.org/10.1109/CDC.2012.6427010

[21] J. Liu and A. R. Teel, "Generalized solutions to hybrid systems with delays," *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pp. 6169–6174, 2012. [Online]. Available: http://dx.doi.org/10.1109/CDC.2012.6425940

[22] V. I. Utkin, J. Guldner, and J. Shi, *Sliding mode control in electromechanical systems*, ser. The Taylor & Francis systems and control book series. London, Philadelphia, PA: Taylor & Francis, 1999.

[23] J. Liu and X. Wang, *Advanced Sliding Mode Control for Mechanical Systems : Design, Analysis and MATLAB Simulation*. Berlin: Springer Berlin, 2012, 11,N15. [Online]. Available: http://opac.inria.fr/record=b1133965

[24] A. Polyakov and L. Fridman, "Stability notions and Lyapunov functions for sliding mode control systems," *J. Franklin Inst.*, vol. 351, no. 4, pp. 1831–1865, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.jfranklin.2014.01.002

[25] R. Khasminskii, *Stochastic stability of differential equations*, 2nd ed., ser. Stochastic Modelling and Applied Probability. Heidelberg: Springer, 2012, vol. 66, with contributions by G. N. Milstein and M. B. Nevelson. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-23280-0

[26] X. Mao and C. Yuan, *Stochastic differential equations with Markovian switching*. London: Imperial College Press, 2006.

[27] J. George, "Robust fault detection and isolation in stochastic systems," *Internat. J. Control*, vol. 85, no. 7, pp. 779–799, 2012. [Online]. Available: http://dx.doi.org/10.1080/00207179.2012.666360

[28] V. I. Utkin, "Sliding mode control design principles and applications to electric drives," *IEEE Transactions on Industrial Electronics*, vol. 40, no. 1, pp. 23–35, 1993.

[29] A. F. Filippov, *Differential equations with discontinuous righthand sides*, ser. Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht, 1988, vol. 18, translated from the Russian. [Online]. Available: http://dx.doi.org/10.1007/978-94-015-7793-9

[30] J. Cortés, "Discontinuous dynamical systems: a tutorial on solutions, nonsmooth analysis, and stability," *IEEE Control Syst. Mag.*, vol. 28, no. 3, pp. 36–73, 2008. [Online]. Available: http://dx.doi.org/10.1109/MCS.2008.919306

[31] E. A. Asarin and R. N. Izmaĭlov, "Determination of the sliding speed on a discontinuity surface," *Avtomat. i Telemekh. (transl.,Automation and Remote Control 50(1989),pp.1181-1185)*, no. 9, pp. 43–48, 1989.

[32] L. Arnold, *Stochastic differential equations: theory and applications*. New York: Wiley-Interscience [John Wiley & Sons], 1974, translated from the German.

[33] K. Ito, "On stochastic differential equations," *Mem. Amer. Math. Soc.*, vol. 1951, no. 4, p. 51, 1951.

[34] R. L. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM J. Control*, vol. 4, pp. 362–371, 1966.

[35] R. F. Bass, *Stochastic processes*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press, 2011, vol. 33.

[36] T. Yamada and S. Watanabe, "On the uniqueness of solutions of stochastic differential equations." *J. Math. Kyoto Univ.*, vol. 11, pp. 155–167, 1971.

[37] A. S. Cherny and H.-J. Engelbert, *Singular stochastic differential equations*, ser. Lecture Notes in Mathematics. Berlin: Springer-Verlag, 2005, vol. 1858.

[38] A. K. Zvonkin, "A transformation of the phase space of a diffusion process that will remove the drift," *Mat. Sb. (N.S.)*, vol. 93(135), pp. 129–149, 152, 1974.

[39] A. J. Veretennikov, "On strong solutions and explicit formulas for solutions of stochastic integral equations," *Math. USSR Sb.*, vol. 39, no. 3, pp. 387–403, 1981. [Online]. Available: http://dx.doi.org/10.1070/SM1981v039n03ABEH001522

[40] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, ser. Applications of Mathematics (New York). Berlin: Springer-Verlag, 1992, vol. 23.

[41] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM Rev.*, vol. 43, no. 3, pp. 525–546 (electronic), 2001. [Online]. Available: http://dx.doi.org/10.1137/S0036144500378302

[42] G. Maruyama, "Continuous Markov processes and stochastic equations," *Rend. Circ. Mat. Palermo (2)*, vol. 4, pp. 48–90, 1955.

[43] C. Yuan and X. Mao, "Convergence of the Euler-Maruyama method for stochastic differential equations with Markovian switching," *Math. Comput. Simulation*, vol. 64, no. 2, pp. 223–235, 2004. [Online]. Available: http://dx.doi.org/10.1016/j.matcom.2003.09.001

[44] Y. Wang and C. Yuan, "Convergence of the Euler-Maruyama method for stochastic differential equations with respect to semimartingales," *Appl. Math. Sci. (Ruse)*, vol. 1, no. 41-44, pp. 2063–2077, 2007.

[45] I. Gyöngy and N. Krylov, "Existence of strong solutions for Itô's stochastic equations via approximations," *Probab. Theory Related Fields*, vol. 105, no. 2, pp. 143–158, 1996. [Online]. Available: http://dx.doi.org/10.1007/BF01203833

[46] K. S. Chan and O. Stramer, "Weak consistency of the Euler method for numerically solving stochastic differential equations with discontinuous coefficients," *Stochastic Process. Appl.*, vol. 76, no. 1, pp. 33–44, 1998. [Online]. Available: http://dx.doi.org/10.1016/S0304-4149(98)00020-9

[47] L. Yan, "The Euler scheme with irregular coefficients," *Ann. Probab.*, vol. 30, no. 3, pp. 1172–1194, 2002. [Online]. Available: http://dx.doi.org/10.1214/aop/1029867124

[48] G. Leobacher and M. Szölgyenyi, "A numerical method for SDEs with discontinuous drift," *BIT*, vol. 56, no. 1, pp. 151–162, 2016. [Online]. Available: http://dx.doi.org/10.1007/s10543-015-0549-x

[49] H.-L. Ngo and D. Taguchi, "Strong convergence for the Euler–Maruyama approximation of stochastic differential equations with discontinuous coefficients," *Statist. Probab. Lett.*, vol. 125, pp. 55–63, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.spl.2017.01.027

[50] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the brownian motion," *Physical Review*, vol. 36, pp. 823–841, 1930.

[51] H. Risken, *The Fokker-Planck equation*, 2nd ed., ser. Springer Series in Synergetics. Berlin: Springer-Verlag, 1989, vol. 18, methods of solution and applications. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-61544-3

[52] D. Sundararajan, *The discrete Fourier transform*. World Scientific Publishing Co., Inc., River Edge, NJ, 2001, theory, algorithms and applications.

[53] R. P. Agarwal and D. O'Regan, *Ordinary and partial differential equations*, ser. Universitext. Springer, New York, 2009, with special functions, Fourier series, and boundary value problems.

References

Part II

Contributions

Paper A

A Simple Stochastic Differential Equation with Discontinuous Drift

Maria Simonsen, John Leth, Henrik Schiøler and Horia Cornean

The paper has been published in the
proceedings of the Third International Workshop on Hybrid Autonomous Systems,
2013.

The layout has been revised, and small editorial changes have been made. Content relevant changes, if any, are marked with explicit footnotes.

Abstract

In this paper we study solutions to stochastic differential equations (SDEs) with discontinuous drift. We apply two approaches: The Euler-Maruyama method and the Fokker-Planck equation and show that a candidate density function based on the Euler-Maruyama method approximates a candidate density function based on the stationary Fokker-Planck equation. Furthermore, we introduce a smooth function which approximates the discontinuous drift and apply the Euler-Maruyama method and the Fokker-Planck equation with this input. The point of departure for this work is a particular SDE with discontinuous drift.

1 Introduction

Since the pioneering work by Einstein on Brownian motion [1], stochastic differential equations (SDEs) have been intensively studied, with the foundation for SDEs developed by Itô and Stratonovich, e.g. see [2, 3] and references therein. Well-developed theories for various sub-disciplines of SDEs have been developed ever since, such as stability- and control theory [4–6]. Lately SDEs have been used in the generalization of hybrid dynamical systems [7–9]. Moreover, today a variety of applications and numerical methods exist for SDEs, see [10, 11] and references therein.

Comment to the above theories is that they are all developed under (weak) regularity conditions on the drift and diffusion coefficients. Such conditions are, of course, necessary in order to develop an applicable/operational theory. However, it is also of interest to study special cases (or classes) where the standard regularity conditions fails but fundamental properties of the SDEs are still valid such as existence and uniqueness of solutions. As an example, we recall that a necessary condition for existence of solutions for an (deterministic) ordinary differential equation (ODE), $\dot{x} = f(x)$, is that the right-hand side of the ODE, that is the vector field f, should be continuous in the state variable x. However, the Cauchy problem $\dot{x} = \mathrm{sgn}(x),\ x(0) = x_0$, with sgn denoting the sgn-function, has a solution for all initial values x_0, despite the fact that sgn is discontinuous at zero (we remark that $\mathrm{sgn}(0) = 0$ by definition, if this was not the case, no solution would exist at $x_0 = 0$). As discontinuous functions often appear in applications, ODEs such as $\dot{x} = \mathrm{sgn}(x)$ should and have been studied, see [12] and references therein.

In this work in progress we initiate a study of SDEs with discontinuous drift. This should be seen as a part of a larger scope with focuses on the intersection between SDEs and switching dynamics, a field only scarcely explored so far, see [13] and references therein for related work. Within this larger scope theoretical questions to be answered are how to define solutions to SDEs with state dependent switching (in particular with discontinuous drift), when solutions cannot be patched together of segments of positive time duration and can any applicable results be obtained by applying the Euler-Maruyama method [14, 15], which is a simple time discrete approximations

technique, to SDEs with drift which do not meet the regularity conditions? Here, we exemplify these problems by studying solution candidates to a particular SDE having discontinuous drift coefficient. More precisely, we consider the SDE

$$dx_t = -k\,\mathrm{sgn}(x_t)dt + dB_t, \tag{A.1}$$

with k a control gain, and ask how a solution can be defined.

The systems treated in [7] are assumed to have non-Zeno execution in finite time. A local behaviour of the Wiener process B_t is that in every time interval of the form $[0, \epsilon)$, with $\epsilon > 0$, B_t has infinitely many zeros, that is, the process which fulfils

$$dx_t = dB_t$$

crosses zero infinitely many times, see [16]. In (A.1) the process x_t is forced to proceed against zero, so we conjecture that the solution to (A.1) (if it exists) also crosses zero infinitely many times in finite time.

We use several methods in the attempt to give a meaningful/operational definition of solutions to (A.1) based on density functions and their probabilistic properties. We start by using the Euler-Maruyama method to approximate numerical solutions to (A.1), which values are presented in histograms for particular time instant. These histograms can be considered as an approximation to the density function for solutions to (A.1). We investigate the influence of the step-size and the control gain in the simulations. Furthermore, from the Euler-Maruyama method we obtain recursively defined density functions which, under stationary conditions, appear to converge to the outcome of the second approach which departures from the Fokker-Planck equation. It is interesting to note that the candidate density function obtained from the Fokker-Planck equation is derived under the assumption of stationarity while the recursive approach has no such assumptions. More precisely, we obtain formulas which strongly indicate that if such a stationary density function exists then it solves the stationary Fokker-Planck equation.

As a third method, we introduce an approximation to the sgn-function and apply both the Euler-Maruyama method and the Fokker-Planck equation to this. A comparison with the stationary density function which solves the Fokker-Planck equation is made.

Finally, we briefly mention one approach which relates to the Euler-Maruyama method. Even though this intuitively should provide some information, so far we have not been able to obtain any meaningful results based on this method. It is included since it is believed that it does in fact carry important information.

It is important to emphasize that the presented material is work in progress and that the heuristic presented here is an initial attempt to define meaningful candidate density functions to solutions to a particular SDE with discontinuous drift. It is clear that for future work the presented material have to be set in a formal mathematical frame including proofs which validate the various procedure used to obtain candidate density functions.

1.1 A Stochastic Differential Equation

A general one dimensional SDE is given by

$$dx_t = b(t, x_t)dt + \sigma(t, x_t)dB_t, \quad x_0 = c \tag{A.2}$$

where $x = x_t$ is an $\mathbb{R}$-valued stochastic process : $[0, T] \to \mathbb{R}$, $b, \sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ are the drift and diffusion coefficient of x, $B = B_t$ is an $\mathbb{R}$-valued Wiener process, and c is a random variable independent of $B_t - B_0$ for $t \geq 0$. On $[0, T]$, existence and uniqueness of a solution x_t, continuous with probability 1, to (A.2) is guaranteed whenever the drift b and diffusion σ are measurable functions satisfying a Lipschitz condition together with a growth bound, both uniformly in t [10, Theorem 5.2.1].

This paper focuses on the special case for (A.2), where $\sigma(t, x_t) = 1$ and $b(t, x_t) = -k\,\mathrm{sgn}(x_t)$ with $k > 0$ a control gain and the sgn-function defined by

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{if} \quad x < 0 \\ 0 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x > 0 \end{cases} .$$

Thus, we consider the SDE

$$dx_t = -k\,\mathrm{sgn}(x_t)dt + dB, \quad x_0 = c \tag{A.3}$$

with c given.

In the next section, the Euler-Maruyama method is applied to approximate solutions to (A.3) and to investigate a theoretical method to obtain candidate density functions for solutions to (A.3).

2 The Euler-Maruyama Method

The Euler-Maruyama method is a simple time discrete approximation technique which is used to approximate solutions to SDEs of the type given in (A.2), by discretizing the time interval $[0, T]$ in steps $0 < t_1 < \cdots < t_n < t_{n+1} \cdots < t_N$ with $N = \left\lceil \frac{T}{h} \right\rceil$, where $h = t_{n+1} - t_n$ is the step-length. Each recursive step is determined via the following method,

$$x_{n+1} = x_n + hb(t_n, x_n) + \sigma(t_n, x_n)W_n \tag{A.4}$$

where $x_{t_i} = x_i$ and $W_n = B_{t_{n+1}} - B_{t_n}$ is i.i.d. normal with mean zero and variance h, which we denote by $W_n \sim N(0, h)$.

Given an initial condition $x_0 = c$, it is possible from (A.4) to approximate a solution to (A.2) by determination of $x_1, x_2, \ldots, x_N$. If the drift and diffusion coefficient in (A.2) are measurable, satisfy a Lipschitz condition and a growth bound, the Euler-Maruyama method guarantees strong convergence to the solution of (A.2), [11, Theorem 9.6.2]. Hence for SDEs with discontinuous drift we can, in general, not expect

the Euler-Maruyama method to produce meaningful results. Nevertheless, we will in the sequel apply this method to the special case (A.3) in order to obtain candidate solutions.

2.1 Analysis of the Deterministic Step

Application of the Euler-Maruyama method to the SDE in (A.3) gives the recursive step

$$x_{n+1} = x_n - hk\,\mathrm{sgn}(x_n) + W_n\ . \tag{A.5}$$

Now, if $x_n > 0$, we have

$$x_{n+1} = x_n - hk + W_n.$$

Since $W_n \sim N(0,h)$, the expectation is that $x_{n+1} \in [x_n - h(k+1), x_n - h(k-1)]$ in most of the simulations. From this, we expect after a finite time $0 < t < \infty$ that there exists $N \in \mathbb{N}$ such that $x_{n+N} \leq 0$. Similar result is obtained if $x_m < 0$, then we expect that there exists $M \in \mathbb{N}$ such that $x_{m+M} \geq 0$. The influence from the control gain k determines how quick the evolution of the sequence $\{x_n\}_{n\geq 0}$ switches around zero. In other words, a big k minimizes the influence of the random variable W_n.

The Euler-Maruyama method is easy to implement in software, so following we have applied Matlab to simulate solutions to (A.3).

2.2 Numerical Solutions to an SDE with Discontinuous Drift

We consider the recursive step in (A.5) and simulate the evolution of the stochastic process. For all simulations the initial condition is chosen to be $x_0 = 0$ and the considered time interval is $[0,T]$ where $T = 1$. The step-length is h such that the number of simulated steps is $N = \left\lceil \frac{T}{h} \right\rceil$. All simulations are repeated 500 times and histograms of the resulting values of x_T are presented.

In Figure A.1, one realization of a solution to the SDE in (A.3) is shown together with the average values of all the 500 simulations in the time interval $[0,T]$. The average of x_t is close to zero for all $0 \leq t \leq T$.

Figure A.2 shows the resulting histogram of x_T including 500 simulations.

In order to investigate the influence of the step-size, Figure A.3 illustrates four different histograms of 500 simulations. Here h is $0.01, 0.001, 0.0001$ and 0.00001 respectively and $k = 1$. It can be seen that the result narrows slightly around zero when h becomes smaller, but changing in the step-size does not immediately give big effect.

In Figure A.4, the control gain is changing, $k = 1,2,3,4$ and $h = 0.001$. Here it is clear that changing k has an influence on the result of x_T. The variance of x_T gets smaller when k increases. This is not surprising since the overall influence of the random variable W_n is decreased when k increases as mentioned previously.

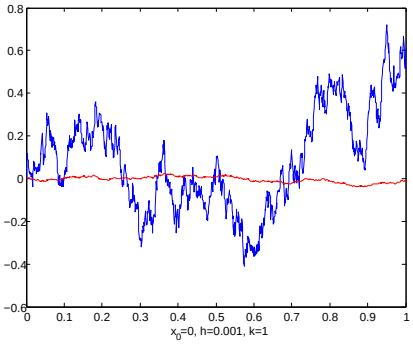

Figure A.1: The blue graph shows one realization out of 500 simulations while the red graph represents the average of x_t for $0 \leq t \leq T$.

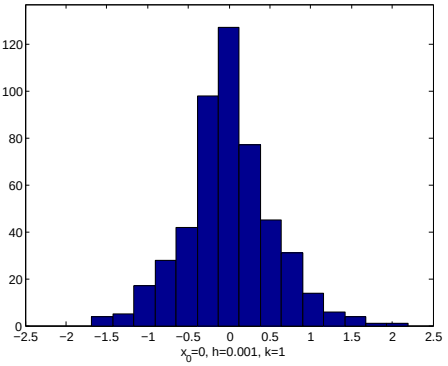

Figure A.2: Resulting histogram of 500 simulations of x_T.

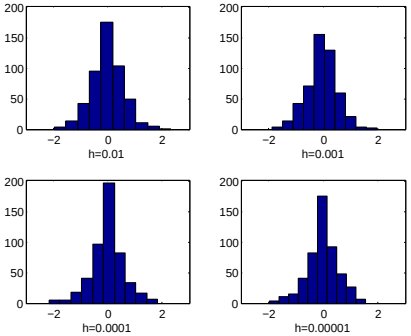

Figure A.3: Histograms of 500 simulations of x_T with different step-size.

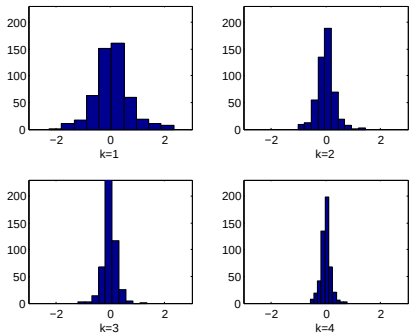

Figure A.4: Histograms of 500 simulations of x_T with different control gains. $h = 0.001$.

2.3 Theoretical Distribution of x_n

In the following, we investigate the distribution of x_n defined by the Euler-Maruyama method.

Consider the recursive determination of x_{n+1} in (A.5). Define an intermediate variable $z_n = x_n + hk\,\mathrm{sgn}(x_n)$ and let f_{z_n} and f_n denote the density functions of z_n and x_n, respectively. Moreover, let $N(0,h)$ denote the density function for W_n. From probability theory we have

$$f_{n+1}(x) = f_{z_n} * N(0,h) \tag{A.6}$$

where $*$ denotes convolution.

Hence, we proceed by studying the density function $f_{z_n}(z)$. Let the distribution function of z_n be denoted by F_{z_n} such that

$$F_{z_n}(z) = P(z_n \leq z) = P(x_n - hk\,\mathrm{sgn}(x_n) \leq z),$$

which can be expressed by

$$\begin{aligned} P(z_n \leq z) &= P(x_n - hk \leq z, x_n > 0) + P(x_n + hk \leq z, x_n < 0) \\ &\quad + P(x_n \leq z, x_n = 0) \\ &= P(x_n \leq z + hk, x_n > 0) + P(x_n \leq z - hk, x_n < 0)\,. \end{aligned}$$

For different values of z, the probability $P(z_n \leq z)$ can be expressed differently. If $z < -hk$

$$P(z_n \leq z) = P(x_n \leq z - hk, x_n < 0) = P(x_n \leq z - hk) = F_n(z - hk)\,.$$

If $z > hk$

$$\begin{aligned} P(z_n \leq z) &= P(x_n \leq z + hk, x_n > 0) + P(x_n \leq z - hk) \\ &= P(x_n \leq z + hk) = F_n(z + hk)\,, \end{aligned}$$

and if $-hk \leq z \leq hk$

$$\begin{aligned} P(z_n \leq z) &= P(x_n \leq z + hk) - P(x_n < 0) + P(x_n \leq z - hk) \\ &= F_n(z + hk) - F_n(0) + F_n(z - hk)\,. \end{aligned}$$

By introducing the indicator function $\mathbb{I}$, the expression of the distribution function of z_n is

$$\begin{aligned} P(z_n \leq z) &= F_n(z - hk)\mathbb{I}_{(-\infty,-hk)}(z) + (F_n(z + hk) - F_n(0) \\ &\quad + F_n(z - hk))\mathbb{I}_{[-hk,hk]}(z) + F_n(z + hk)\mathbb{I}_{(hk,\infty)}(z) \\ &= F_n(z - hk)\mathbb{I}_{(-\infty,hk]}(z) + F_n(z + hk)\mathbb{I}_{[-hk,\infty)}(z) - F_n(0)\mathbb{I}_{[-hk,hk]}(z)\,. \end{aligned}$$

By differentiating with respect to z, the density function of z_n is

$$\begin{aligned}\frac{\partial}{\partial z}F_{z_n} &= f_n(z-hk)\mathbb{I}_{(-\infty,hk]}(z) - F_n(z-hk)\delta(z-hk) + f_n(z+hk)\mathbb{I}_{[-hk,\infty)}(z)\\ &\quad +F_n(z+hk)\delta(z+hk) - F_n(0)(\delta(z+hk)-\delta(z-hk))\\ &= f_n(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f_n(z+hk)\mathbb{I}_{(-hk,\infty)}(z)\\ &\quad +\delta(z+hk)(F_n(z+hk)-F_n(0)) + \delta(z-hk)(F_n(0)-F_n(z-hk))\\ &= f_n(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f_n(z+hk)\mathbb{I}_{[-hk,\infty)}(z)\ .\end{aligned}$$

Therefore

$$f_{z_n}(z) = f_n(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f_n(z+hk)\mathbb{I}_{[-hk,\infty)}(z)\ . \tag{A.7}$$

By substituting the above into (A.6), the density function f_{n+1} is found from the density function f_n. In the following section, the solution to (A.7) is investigated numerically.

2.4 Recursive Developing of the Density Function in Matlab

The recursive density function is given by

$$f_{n+1}(x) = \Big(f_n(x+kh)\mathbb{I}_{(-\infty,hk]} + f_n(x-hk)\mathbb{I}_{[-hk,\infty)}\Big) * N(0,h)\ . \tag{A.8}$$

Following, we apply Matlab to investigate the evolution of the function $f_{n+1}(x)$ for increasing n. Assume that the density function for the initial condition $x_0 = c$ is normal distributed with mean zero and variance h. The Euler-Maruyama method is expected to converge to a stochastic process (or a distribution of a stochastic process) when $h \to 0$. (Under certain regularity conditions, so actually we cannot expect it here but only conjecture.) We hope that the developing of the recursive density functions in (A.8) will reach stationary condition for $n \to \infty$. For this reason, the number n of simulations is chosen to depend on the step size, such that $n = \left\lceil \frac{1}{h^{1+\alpha}} \right\rceil$, where $\alpha > 0$. This ensures that both convergence criteria are fulfilled.

Equation (A.8) is simulated in Matlab for $h = 0.01, \alpha = 0.5, k = 1$ such that $n = 1000$, the result is shown in Figure A.5. At the end of section 3, a comparison between the convergence of the recursive density function and the result obtained there is presented.

In the following, we continue the study of (A.7) under stationary assumptions.

2.5 Stationary State

In the sequel, we continue the study of the recursive density function under the assumptions that it is possible to reach stationary state in (A.8) for $n \to \infty$, such that

$$f_{z_n}(z) = f(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f(z+hk)\mathbb{I}_{(-hk,\infty)}(z)\ ,$$

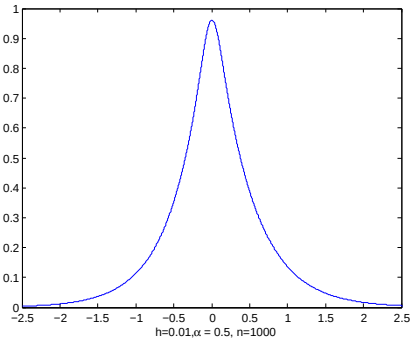

Figure A.5: Recursive density function, f_{1000}.

where $f(x)$ is the stationary density function of the recursive development of $f_n(x)$. In this case we define an operator H_h taking f_n to f_{n+1} by

$$\begin{aligned} H_h[f](x) &= \int_{\mathbb{R}} \Big(f(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f(z+hk)\mathbb{I}_{(-hk,\infty)}(z) \Big) \\ &\quad \cdot \frac{1}{\sqrt{2\pi h}} \exp\left(\frac{-(x-z)^2}{2h} \right) dz \, . \end{aligned}$$

By taking the derivative with respect to h and then the limit $h \to 0$, we obtain (see the Appendix),

$$\lim_{h\to 0} \frac{\partial}{\partial h} H_h[f](x) = \begin{cases} -kf'(x) + \frac{1}{2}f''(x) & \text{for} \quad x < 0 \\ \psi(x) & \text{for} \quad x = 0 \\ kf'(x) + \frac{1}{2}f''(x) & \text{for} \quad x > 0 \end{cases} . \tag{A.9}$$

The case $x = 0$ is not important for the sequel development, hence we leave $\psi(x)$ unspecified. Note that the right hand side of the first and last case in (A.9) have similarity with the stationary Fokker-Planck equation.

An approximation of the operator H_h is

$$\begin{aligned} H_h[f](x) &\approx f(x) + h(k\,\mathrm{sgn}(x)f'(x) + \frac{1}{2}f''(x)) \\ &= f(x) + hG[f](x) \, . \end{aligned} \tag{A.10}$$

For a fixed h define f_h to be the stationary density function and assume that $\lim_{h\to 0} f_h$ exists, say f_0 and that $H_h(f_h) = f_h$. Hence, if we disregard the approximation in (A.10), we look for a function f_h such that $G[f_h] = 0$, which by (A.9) means that f_h is a stationary solution to the Fokker-Planck equation. We conjecture that this heuristic will constitute the main ideas in the proof that the stationary distribution f_0 $(= \lim_{h\to 0} f_h)$ of the Euler-Maruyama simulation converges to the Fokker-Planck equation.

3 The Fokker-Planck Equation

Based on the preceding work, this section introduces the one-dimensional Fokker-Planck equation and applies this to determine a density function of x_t for fixed t.

Let x_t be a solution to (A.3). For a fixed $t \in [0,T]$ let $\phi(x,t)$ be the density function of x_t, such that

$$\int_{\mathbb{R}} \phi(x,t)dx = 1 \tag{A.11}$$

with initial condition

$$\lim_{t\to 0} \phi(x,t) = \delta(x - x_0)$$

where δ denotes the Dirac delta function. If the drift and diffusion coefficient are moderately smooth functions, then the density function $\phi(x,t)$ satisfies the Fokker-Plank equation [16, 17]. For (A.3) this means

$$\frac{\partial}{\partial t}\phi(x,t) = \frac{\partial}{\partial x}k\,\mathrm{sgn}(x)\phi(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi(x,t)\ . \tag{A.12}$$

As mentioned previously, the drift coefficient $-k\,\mathrm{sgn}(x)$ is not a smooth function and for this reason there is no guarantee that (A.12) is valid.

3.1 Solution in Two Domains

In order to avoid the discontinuous challenges by the sgn-function, we consider the Fokker-Planck equation (A.12) in the domains $(-\infty,0)$ and $(0,\infty)$. This gives

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(x,t) &= \frac{\partial}{\partial x}k\phi(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi(x,t) \quad \text{for } x > 0 \\
\frac{\partial}{\partial t}\phi(x,t) &= -\frac{\partial}{\partial x}k\phi(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi(x,t) \quad \text{for } x < 0\ .
\end{aligned}$$

Assume that the density function can reach stationarity, then

$$\begin{aligned}
0 &= \frac{\partial}{\partial x}k\phi(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi(x,t) \quad \text{for } x > 0 \\
0 &= -\frac{\partial}{\partial x}k\phi(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi(x,t) \quad \text{for } x < 0\ ,
\end{aligned}$$

which are two ODEs. The characteristic equations are

$$\begin{aligned}
ks + \frac{1}{2}s^2 &= 0 \quad \text{for } x > 0 \\
-ks + \frac{1}{2}s^2 &= 0 \quad \text{for } x < 0\ ,
\end{aligned}$$

which have the roots

$$\begin{aligned}
s = 0, \qquad s = -2k \quad &\text{for } x > 0 \\
s = 0, \qquad s = 2k \quad &\text{for } x < 0\ ,
\end{aligned}$$

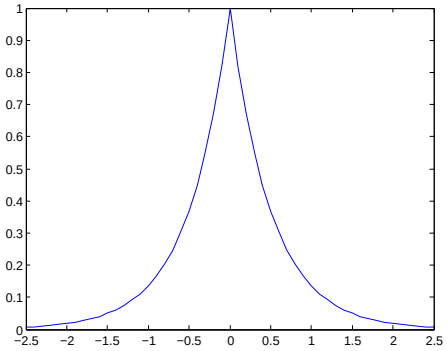

Figure A.6: Density function $\phi(x,t)$ for x_t for fixed $t \in [0,T]$.

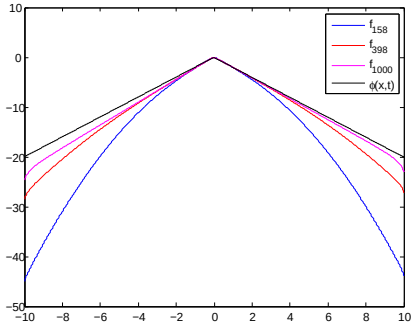

Figure A.7: Logarithmic stationary density function and logarithmic recursive density function of $f_{158}, f_{398}, f_{1000}$.

such that

$$\begin{aligned} \phi_+(x) &= c_1 \exp(0x) + d_1 \exp(-2kx) \quad \text{for } x > 0 \\ \phi_-(x) &= c_2 \exp(0x) + d_2 \exp(2kx) \quad \text{for } x < 0 \end{aligned}$$

are solutions of (A.12) in the respective domain. We discard the constant term and, due to symmetry around $x = 0$, it can be expected that $d_1 = d_2 := d$. The boundary constraint in equation (A.11) gives

$$d = \frac{1}{2\int_0^\infty \exp(-2kx)} = k\,.$$

Furthermore, from the above we assume that $\phi(x,t) = k$ for $x = 0$. This gives the density function

$$\phi(x,t) = \begin{cases} k\exp(-2kx) & \text{for } x > 0 \\ k\exp(2kx) & \text{for } x < 0 \\ k & \text{for } x = 0 \end{cases} \tag{A.13}$$

which is illustrated in Figure A.6 for $k = 1$. The density function $\phi(x,t)$ is the Laplace distribution with location parameter zero and scale parameter $\frac{1}{2k}$.

The density function in (A.13) is compared with the recursive density functions developed in section 2.4 by considering the logarithmic of the density functions. The recursive cases with $n = 158, 398, 1000$ are shown in Figure A.7 together with the logarithmic stationary density function $\log(\phi(x,t))$. For increasing n, the logarithmic recursive density function gets closer to the logarithmic stationary density function, as we hope to observe.

Due to the lack of continuity in all the previous investigations, in the following section we construct a smooth function which approximates the sgn-function and investigate what results the applied methods provide with this smooth function as input.

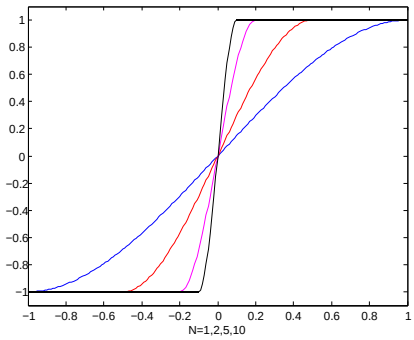

Figure A.8: Approximation of the sgn-function.

4 Approximation of sign-Function

As mentioned, the problem with the sgn-function is that it is not smooth. For this reason we construct a function $f_N(x)$ which has pointwise convergence to the sgn-function as $N \to \infty$.

$$f_N(x) = \begin{cases} 1 & \text{for } x > \frac{1}{N} \\ -\frac{N^3}{2}x^3 + \frac{3N}{2}x & \text{for } -\frac{1}{N} \leq x \leq \frac{1}{N} \\ -1 & \text{for } x < -\frac{1}{N} \end{cases} . \tag{A.14}$$

In Figure A.8, the functions $f_1(x), f_2(x), f_5(x)$ and $f_{10}(x)$ are shown. In the following, we will apply the Euler-Maruyama method to this function.

4.1 Euler-Maruyama Method Applied to the Approximated sign-Function

It is proved that the Euler-Maruyama method exhibits convergence to the solution of the SDE if the drift and diffusion coefficient satisfy certain regularity conditions. Since $f_N(x)$ from (A.14) is a smooth function, we apply the Euler-Maruyama method with this input instead of the sgn-function. The same procedure as in section 2.2 is applied, such that

$$x_{j+1} = x_j - hkf_N(x_j) + W_j . \tag{A.15}$$

Since the slope of $f_N(x)$ increases significantly with the increase of N, the step-size is chosen to be dependent on N, such that $h = \frac{0.001}{N}$. We hope to avoid inaccurate simulations due to the step-size with this method. (If $x_i \in [-\frac{1}{N}, \frac{1}{N}]$ for $t_i \in [0, T]$ and N is big, then the value of $f_N(x_i)$ is dominating in the determination of x_{i+1} such that $x_i << x_{i+1}$ or $x_i >> x_{i+1}$.) As before, $k = 1$. In Figure A.9 3 realizations of x_t are shown, when $f_4(x)$ is used to approximate $\text{sgn}(x)$. In Figure A.10 histograms of 500 simulations are presented based on different approximations of the sgn-function.

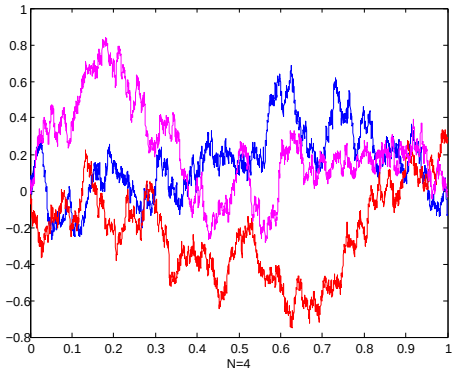

Figure A.9: Realizations of x_t to an SDE with drift $b(t, x_t) = -kf_4(x_t)$.

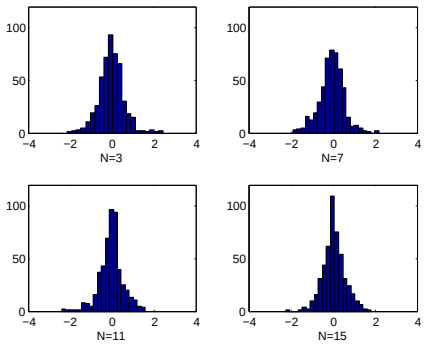

Figure A.10: Histogram of 500 simulations of x_T where different approximations to the sgn-function are used.

In the following section, the smooth function $f_N(x)$ is implemented in the Fokker-Planck equation.

4.2 Solution to the Fokker-Planck Equation with Approximated sign-Function

Previously, the Fokker-Planck equation is applied to an SDE where the drift is discontinuous due to the sgn-function. As mentioned, we cannot expect this to be meaningful. Following, the smooth function $f_N(x)$ is used as a substitute for the sgn-function, such that we consider the SDE

$$dx_t = -kf_N(x_t)dt + dB_t \, . \tag{A.16}$$

We investigate if it is possible to determine a density function which is a solution to the Fokker-Planck equation with $-kf_N(x)$ as the drift.

With the same technique as previously, we consider the Fokker-Planck equation in three domains. First, the domain $[-\frac{1}{N}, \frac{1}{N}]$ is considered with $f_N(x)$ as an approximation of $\operatorname{sgn}(x)$,

$$\frac{\partial}{\partial t}\phi_N(x,t) = \frac{\partial}{\partial x}k\left(-\frac{N^3}{2}x^3 + \frac{3N}{2}x\right)\phi_N(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\phi_N(x,t) \, . \tag{A.17}$$

Under stationary assumptions, the function $\phi_N(x,t) = \phi_0 \exp\left(\frac{N^3}{4}kx^4 - \frac{3N}{2}kx^2\right)$ fulfils (A.17). By including the domains $(-\infty, -\frac{1}{N})$ and $(\frac{1}{N}, \infty)$, the density function for solutions to (A.16) becomes

$$\phi_N(x,t) = \begin{cases} d\exp(2kx) & \text{for } x < -\frac{1}{N} \\ \phi_0 \exp\left(\frac{N^3}{4}kx^4 - \frac{3N}{2}kx^2\right) & \text{for } -\frac{1}{N} \leq x \leq \frac{1}{N} \\ d\exp(-2kx) & \text{for } x > \frac{1}{N} \end{cases} \, ,$$

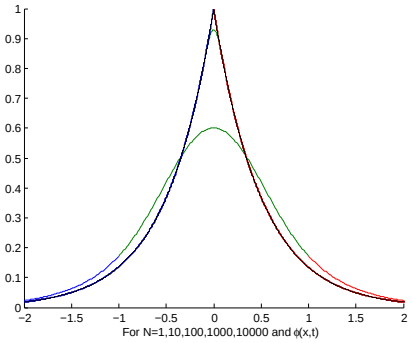

Figure A.11: Density functions for solutions to (A.16) together with $\phi(x,t)$.

with constraints

$$\begin{aligned} 1 &= \int_{-\infty}^{-\frac{1}{N}} d\exp(2kx)dx + \int_{\frac{1}{N}}^{\infty} d\exp(-2kx)dx \\ &\quad + \int_{-\frac{1}{N}}^{\frac{1}{N}} \phi_0 \exp\left(k\left(\frac{N^3}{4}x^4 - \frac{3N}{2}x^2\right)\right)dx \end{aligned}$$

and

$$\lim_{x\to\pm\frac{1}{N}+} \phi_N(x,t) = \lim_{x\to\pm\frac{1}{N}-} \phi_N(x,t)\,. \tag{A.18}$$

From (A.18)

$$d = \phi_0 \exp\left(\frac{3k}{4N}\right),$$

so the normalization constant becomes

$$\phi_0 = \frac{1}{2\int_0^{\frac{1}{N}} \exp\left(-\frac{3kN}{2}x^2 + \frac{kN^3}{4}x^4\right)dx + \frac{1}{k}\exp\left(-\frac{5k}{4N}\right)}$$

and

$$\phi_N(x,t) = \begin{cases} \phi_0 \exp\left(\frac{3k}{4N}\right)\exp(2kx) & \text{for } x < -\frac{1}{N} \\ \phi_0 \exp\left(\frac{N^3}{4}kx^4 - \frac{3N}{2}kx^2\right) & \text{for } -\frac{1}{N} \le x \le \frac{1}{N} \\ \phi_0 \exp\left(\frac{3k}{4N}\right)\exp(-2kx) & \text{for } x > \frac{1}{N} \end{cases}.$$

This density function is simulated in Matlab with ϕ_0 numerically calculated. In Figure A.11 results are shown for $N = 1, 10, 100, 1000, 10000$ together with the stationary density function from the Fokker-Planck equation $\phi(x,t)$. Note that ϕ_{100}, ϕ_{1000} and ϕ_{10000} are not visible. The simulations show that $\phi_N(x,t) \to \phi(x,t)$ as expected.

5 Fourier Transformation

In section 2.3 we investigate a candidate density function for the intermediate variable z_n in order to determine a density function for the stochastic process produced by Euler-Maruyama method. Another method is to make a Fourier transformation of the density function for z_n before the convolution of the density functions is done. The Fourier transformation of (A.7) is given by

$$\hat{F}(\omega) \quad = \quad \int_{\mathbb{R}} \left[f_n(z-hk)\mathbb{I}_{(-\infty,hk]} + f_n(z+hk)\mathbb{I}_{(-hk,\infty)} \right] \exp(-j\omega z)dz \; .$$

By expansion and including of the Laplace transformation, see the Appendix, the following relation appears:

$$\hat{F}(\omega) = \cos(\omega hk)\,\mathrm{Re}\left(\hat{F}(\omega)\right) - 2\sin(\omega hk)\,\mathrm{Im}\left(\tilde{F}(\omega)\right)$$

where $\tilde{F}(\omega)$ is the Laplace transformation. For the moment we have no further interpretation of how this can help the developing of the recursive density function.

6 Discussion

This paper presents initial studies of SDEs with discontinuous drift. Theoretical and numerical approaches are applied to a particular SDE in order to investigate meaningful results in the form of density functions.

The candidate recursive density functions developed from the Euler-Maruyama method have the tendency to approximate the density function of the stationary Fokker-Planck equation, which strengthens the conjecture that these density functions actually exist. Furthermore, the fact that the Laplace distribution is a solution to the stationary Fokker-Planck equation with discontinuous drift supports the assumption that it is possible to give a meaningful answering to questions about definition of solutions to discontinuous SDEs and their probabilistic properties. However, we have not completed these processes, yet.

All in all, the different approaches indicate a connection between the candidate density function for the solution to the particular SDE with discontinuous drift and the Fokker-Planck equation. Though, the stationary assumptions have extensive impact on this result. Without this condition, the solvability of the Fokker-Planck equation would decrease significantly.

A potential object for future studies is to formalize the heuristic presentation in this paper to obtain operational definition and regular results connected to classes of SDEs with discontinuous drift.

Appendix

Approximation of the operator H_h

In section 2.5 the operator H_h was defined by

$$\begin{aligned} H_h[f](x) &= \int_{\mathbb{R}} \Big(f(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f(z+hk)\mathbb{I}_{(-hk,\infty)}(z) \Big) \\ &\quad \cdot \frac{1}{\sqrt{2\pi h}} \exp\left(\frac{-(x-z)^2}{2h} \right) dz \ . \end{aligned}$$

Following, the notation $N_h(z) = \frac{1}{\sqrt{2\pi h}} \exp\left(\frac{-z^2}{2h}\right)$ is used and we consider the generator $\frac{\partial}{\partial h} H_h$.

$$\begin{aligned} &\frac{\partial}{\partial h} H_h[f](x) \\ &= \int_{\mathbb{R}} \Big(-kf'(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + kf'(z+hk)\mathbb{I}_{[-hk,\infty)}(z) \Big) \cdot N_h(x-z)dz \\ &\quad + \int_{\mathbb{R}} \left(-kf(z-hk)\delta(z-hk) + kf(z+hk)\delta(z+hk) \right) \cdot N_h(x-z)dz \\ &\quad + \int_{\mathbb{R}} \Big(f(z-hk)\mathbb{I}_{(-\infty,hk]}(z) + f(z+hk)\mathbb{I}_{[-hk,\infty)}(z) \Big) \frac{\partial}{\partial h}(N_h(x-z))dz \ . \end{aligned} \tag{A.19}$$

The second integral term above gives

$$\begin{aligned} &\int_{\mathbb{R}} \left(-kf(z-hk)\delta(z-hk) + kf(z+hk)\delta(z+hk) \right) \cdot N_h(x-z)dz \\ = \ & kf(0) \left(N_h(x+hk) - N_h(x-hk) \right) \ . \end{aligned}$$

The derivative of the last integral term in (A.19) gives

$$\frac{\partial}{\partial h}(N_h(x-z)) = \frac{1}{\sqrt{2\pi h}} \left(\frac{(x-z)^2}{2h^2} - \frac{1}{2h} \right) \exp\left(\frac{-(x-z)^2}{2h} \right) \ .$$

Consider the derivative with respect to z,

$$\frac{\partial}{\partial z}(N_h(x-z)) = \frac{1}{\sqrt{2\pi h}} \frac{-(x-z)}{h} \exp\left(\frac{-(x-z)^2}{2h} \right) \ ,$$

and the second derivative with respect to z,

$$\frac{\partial^2}{\partial z^2}(N_h(x-z)) = \frac{1}{\sqrt{2\pi h}} \left(\frac{(x-z)^2}{h^2} + \frac{1}{h} \right) \exp\left(\frac{-(x-z)^2}{2h} \right) \ .$$

Thus, for small h

$$\frac{\partial}{\partial h}(N_h(x-z)) \approx \frac{1}{2} \frac{\partial^2}{\partial z^2}(N_h(x-z)) \ .$$

Therefore, the last integral term in $\frac{\partial}{\partial h}H_h[f](x)$ is approximately

$$\begin{aligned}
& \frac{1}{2}\int_{\mathbb{R}}\Big(f(z-hk)\mathbb{I}_{(-\infty,hk]}(z)+f(z+hk)\mathbb{I}_{[-hk,\infty)}(z)\Big)\frac{\partial^2}{\partial z^2}(N_h(x-z))dz \\
= \;& \frac{1}{2}\int_{-\infty}^{hk}f(z-hk)\frac{\partial^2}{\partial z^2}(N_h(x-z))dz+\frac{1}{2}\int_{-hk}^{\infty}f(z+hk)\frac{\partial^2}{\partial z^2}(N_h(x-z))dz \\
= \;& \frac{1}{2}\Big([f(z-hk)\frac{\partial}{\partial z}(N_h(x-z))]_{-\infty}^{hk}-[f'(z-hk)N_h(x-z)]_{-\infty}^{hk} \\
& +\int_{-\infty}^{hk}f''(z-hk)N_h(x-z)dz\Big)+\frac{1}{2}\Big([f(z+hk)\frac{\partial}{\partial z}(N_h(x-z))]_{-hk}^{\infty} \\
& -[f'(z+hk)N_h(x-z)]_{-hk}^{\infty}+\int_{-hk}^{\infty}f''(z+hk)N_h(x-z)dz\Big)\,.
\end{aligned}$$

The limit of $\frac{\partial}{\partial h}H_h$ for $h\to 0$ is then

$$\lim_{h\to 0}\frac{\partial}{\partial h}H_h[f](x)=\begin{cases}-kf'(x)+\frac{1}{2}f''(x) & \text{for} \quad x<0\\ \psi(x) & \text{for} \quad x=0\\ kf'(x)+\frac{1}{2}f''(x) & \text{for} \quad x>0\end{cases}$$

where $\psi(x)$ is unknown.

Fourier Transformation Expansion

In section 5 the Fourier transformation of f_{z_n} is presented. Below follows the expansion.

$$\begin{aligned}
\hat{F}(\omega) &= \int_{\mathbb{R}}\Big[f_n(z-hk)\mathbb{I}_{(-\infty,hk]}+f_n(z+hk)\mathbb{I}_{(-hk,\infty)}\Big]\exp(-j\omega z)dz \\
&= \int_{-\infty}^{hk}f_n(z-hk)\exp(-j\omega z)dz+\int_{-hk}^{\infty}f_n(z+hk)\exp(-j\omega z)dz\,.
\end{aligned}$$

By changing variable $x=z-hk$ in the first integral and $x=z+hk$ in the second integral,

$$\begin{aligned}
\hat{F}(\omega) &= \int_{-\infty}^{0}f_n(x)\exp(-j\omega(x+hk))dx+\int_{0}^{\infty}f_n(x)\exp(-j\omega(x-hk))dx \\
&= \exp(-j\omega hk))\int_{-\infty}^{0}f_n(x)\exp(-j\omega x)dx \\
&\quad +\exp(j\omega hk)\int_{0}^{\infty}f_n(x)\exp(-j\omega x)dx \\
&= (\cos(\omega hk)-j\sin(\omega hk))\int_{-\infty}^{0}f_n(x)\exp(-j\omega x)dx \\
&\quad +(\cos(\omega hk)+j\sin(\omega hk))\int_{0}^{\infty}f_n(x)\exp(-j\omega x)dx\,. \qquad \text{(A.20)}
\end{aligned}$$

Based on the uncertainty around zero in the recursive determination of x_{n+1} in (A.4), we find it reasonable to assume that f_n is an even function. Fourier transformations of even functions give zero in the imaginary part, so only the real part of (A.20) is considered.

$$\begin{aligned}
\hat{F}(\omega) &= \cos(\omega hk)\,\mathrm{Re}\left(\int_{-\infty}^{0} f_n(x)\exp(-j\omega x)dx\right) \\
&\quad + \sin(\omega hk)\,\mathrm{Im}\left(\int_{-\infty}^{0} f_n(x)\exp(-j\omega x)dx\right) \\
&\quad + \cos(\omega hk)\,\mathrm{Re}\left(\int_{0}^{\infty} f_n(x)\exp(-j\omega x)dx\right) \\
&\quad - \sin(\omega hk)\,\mathrm{Im}\left(\int_{0}^{\infty} f_n(x)\exp(-j\omega x)dx\right) \\
&= \cos(\omega hk)\int_{-\infty}^{0} f_n(x)\cos(\omega x)dx - \sin(\omega hk)\int_{-\infty}^{0} f_n(x)\sin(\omega x)dx \\
&\quad + \cos(\omega hk)\int_{0}^{\infty} f_n(x)\cos(\omega x)dx + \sin(\omega hk)\int_{0}^{\infty} f_n(x)\sin(\omega x)dx \\
&= \cos(\omega hk)\int_{-\infty}^{\infty} f_n(x)\cos(\omega x)dx + 2\sin(\omega hk)\int_{0}^{\infty} f_n(x)\sin(\omega x)dx \\
&= \cos(\omega hk)\,\mathrm{Re}\left(\hat{F}(\omega)\right) + 2\sin(\omega hk)\int_{0}^{\infty} f_n(x)\sin(\omega x)dx\,. \qquad \text{(A.21)}
\end{aligned}$$

Consider the Laplace transformation of $f_n(x)$,

$$\tilde{F}(s) = \mathcal{L}\{f_n(x)\} = \int_{0}^{\infty} f_n(x)\exp(-sx)dx\,. \qquad \text{(A.22)}$$

The imaginary part of (A.22) is

$$\mathrm{Im}(\tilde{F}(s)) = -\int_{0}^{\infty} f_n(x)\sin(sx)dx\,.$$

From this (A.21) can be expressed by

$$\hat{F}(\omega) = \cos(\omega hk)\,\mathrm{Re}\left(\hat{F}(\omega)\right) - 2\sin(\omega hk)\,\mathrm{Im}\left(\tilde{F}(\omega)\right)$$

where $\tilde{F}(\omega)$ is the Laplace transformation.

References

[1] A. Einstein, "Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen," *Annalen der Physik*, vol. 17, no. 8, pp. 549–560, 1905.

[2] K. Ito, "On stochastic differential equations," *Mem. Amer. Math. Soc.*, vol. 1951, no. 4, p. 51, 1951.

[3] R. L. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM J. Control*, vol. 4, pp. 362–371, 1966.

[4] R. Khasminskii, *Stochastic stability of differential equations*, 2nd ed., ser. Stochastic Modelling and Applied Probability. Heidelberg: Springer, 2012, vol. 66, with contributions by G. N. Milstein and M. B. Nevelson. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-23280-0

[5] H. J. Kushner, *Stochastic stability and control*, ser. Mathematics in Science and Engineering, Vol. 33. New York: Academic Press, 1967.

[6] H. Kushner, *Introduction to stochastic control*. Rinehart and Winston, Inc., New York: Holt, 1971.

[7] M. L. Bujorianu and J. Lygeros, "Toward a general theory of stochastic hybrid systems," in *Stochastic hybrid systems*, ser. Lecture Notes in Control and Inform. Sci. Berlin: Springer, 2006, vol. 337, pp. 3–30.

[8] J. Hu, J. Lygeros, and S. Sastry, "Towards a theory of stochastic hybrid systems," *N. Lynch and B. Krogh (Eds.): HSCC 2000, LNCS 1790*, pp. 160–173, 2000.

[9] X. Mao and C. Yuan, *Stochastic differential equations with Markovian switching*. London: Imperial College Press, 2006.

[10] B. Øksendal, *Stochastic differential equations*, 6th ed., ser. Universitext. Berlin: Springer-Verlag, 2003, an introduction with applications. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-14394-6

[11] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, ser. Applications of Mathematics (New York). Berlin: Springer-Verlag, 1992, vol. 23.

[12] A. F. Filippov, *Differential equations with discontinuous righthand sides*, ser. Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht, 1988, vol. 18, translated from the Russian. [Online]. Available: http://dx.doi.org/10.1007/978-94-015-7793-9

[13] C. Le Bris and P.-L. Lions, "Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients," *Comm. Partial Differential Equations*, vol. 33, no. 7-9, pp. 1272–1317, 2008. [Online]. Available: http://dx.doi.org/10.1080/03605300801970952

[14] G. Maruyama, "Continuous Markov processes and stochastic equations," *Rend. Circ. Mat. Palermo (2)*, vol. 4, pp. 48–90, 1955.

[15] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM Rev.*, vol. 43, no. 3, pp. 525–546 (electronic), 2001. [Online]. Available: http://dx.doi.org/10.1137/S0036144500378302

[16] L. Arnold, *Stochastic differential equations: theory and applications*. New York: Wiley-Interscience [John Wiley & Sons], 1974, translated from the German.

[17] H. Risken, *The Fokker-Planck equation*, 2nd ed., ser. Springer Series in Synergetics. Berlin: Springer-Verlag, 1989, vol. 18, methods of solution and applications. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-61544-3

References

Paper B

A Convergence Result for the Euler-Maruyama Method for a Simple Stochastic Differential Equation with Discontinuous Drift

Maria Simonsen, Henrik Schiøler, John Leth and Horia Cornean

The paper has been published in the
proceedings of 2014 American Control Conference, 2014.

The layout has been revised, and small editorial changes have been made. Content relevant changes, if any, are marked with explicit footnotes.

Abstract

The Euler-Maruyama method is applied to a simple stochastic differential equation (SDE) with discontinuous drift. Convergence aspects are investigated in the case where the Euler-Maruyama method is simulated in dyadic points. A strong rate of convergence is presented for the numerical simulations and it is shown that the produced sequences converge almost surely. This is an improvement of the general result for SDEs with discontinuous drift, i.e. that the Euler-Maruyama approximations converge in probability to a strong solution of the SDE. A numerical example is presented together with a confidence interval for the numerical solutions.

1 Introduction

Stochastic differential equations (SDEs) are used to represent real-world phenomena through mathematic models. The theory of SDEs have been intensively studied with the foundation for SDEs developed by Itô and Stratonovich, e.g. see [1, 2] and references therein. This paper focuses on a simple SDE from a class of SDEs, which are often implemented in models of switched systems. Switched systems arise in many systems for automatic control. They can be found in relatively simple systems such as a compressor in a refrigerator, which keeps the temperature stable, in an automatic gearbox, or in complex systems such as the total control over a flying aircraft.

In modelling, one often is interested in the particular solutions to the model/system in order to be able to predict the future behaviour of the system and/or to be able to control parameters which influence it. Solutions to SDEs can be categorized in two primary groups: strong solutions and weak solutions. Roughly speaking, a weak solution is a solution in probability while a strong solution makes the SDE hold almost surely with a given initial condition. We refer to [3] for further information of forms of solutions to general SDEs.

Solutions to SDEs have been studied for many years. It is not obvious that a solution exists to a given SDE. In 1974 A. K. Zvonkin presented a paper which made a transformation of an SDE with a drift coefficient to an SDE without a drift coefficient [4]. This was an important result since in certain cases processes which only depend on the diffusion coefficient can be solved easier. Hence, this approach made it possible to construct strong solutions to SDEs with irregular drift coefficient. For this transformation to bring a valid strong solution, it is required that the drift coefficient is bounded, measurable and fulfils a Dini condition. Later, A. J. Veretennikov extended this result and showed that sufficient conditions for a strong solution are: the diffusion coefficient is equal the identity matrix and the drift coefficient is a bounded measurable function [5]. In this case the solution is strongly unique.

Since many SDEs do not have closed formed solutions, a common way to obtain information about solutions is by numerical approximation. The Euler-Maruyama method [6] is a stochastic extension of the well-known Euler method for numerical

solution of differential equations. In general, this method guarantees weak and strong convergence to the stochastic process if the drift and diffusion coefficient satisfy appropriate regularity conditions [7]. For certain classes of SDEs, convergence properties for the Euler-Maruyama method have been proven, among others cf. [8–10].

The SDE studied in this paper is an example from a class of SDEs with discontinuous drift coefficient. In 1996 I. Gyöngy and N. Krylov studied such an example and proved that the Euler-Maruyama approximation converges uniformly on bounded intervals, in probability, to a stochastic process. This process was shown to be the strong solution to the studied SDE [11]. They constructed the strong solution by smooth approximations of the coefficients and the initial condition.

In [12] weak convergence of the Euler-Maruyama method is proven for SDEs with discontinuous coefficients. There it is required that the drift and diffusion coefficient are continuous on each element of a family of pairwise disjoint sets, which intersection is the whole space. Moreover, it is assumed that a unique solution exists. In that sense, the outcome of [12] generalizes the weak convergence result for the Euler-Maruyama method when the coefficients are continuous. Weak convergence of the Euler-Maruyama method was also provided by L. Yan where the coefficients are discontinuous on a set with Lebesgue measure zero, and have at most linear growth [13]. He also presented a strong rate of convergence for Euler-Maruyama method to a unique weak solution in the one-dimensional case, under the assumption that the coefficients are Hölder continuous and the drift is Lipschitz. In particular, the main contribution is that he did not assume the Lipschitz condition for the diffusion coefficient. The rate of convergence depends on smooth conditions for the coefficients.

Also Kohatsu-Higa, Lejay and Yasuda have given a result for the weak rate of convergence of the Euler-Maruyama method applied to a multi-dimensional SDE with discontinuous drift. Instead of considering the original discontinuous drift coefficient, they implemented an approximation to the drift in the Euler-Maruyama method and provided a rate of weak convergence to the solution of the original SDE [14]. They also presented some examples in their paper to give a numerical rate of convergence. One of these examples is almost identical with the SDE studied in this paper. In this case the weak convergence rate of the Euler-Maruyama method has order one, which is identical with the order of weak convergence for the Euler-Maruyama method when it is applied to SDEs with sufficiently regular coefficients. In this paper we provide a strong rate of convergence for the Euler-Maruyama method in probability.

We apply the Euler-Maruyama method to a simple SDE with discontinuous drift

$$dx_t = -k\,\mathrm{sgn}(x_t)dt + dW_t, \tag{B.1}$$

with k a control gain and prove convergence of the Euler-Maruyama method with probability one. From [5] we already know that the strong solution to (B.1) exists and is strongly unique, since the drift coefficient is bounded and Lebesgue measurable and the diffusion coefficient is one. Hence, by [5], it follows that the numerical solutions constructed from the Euler-Maruyama method converge to the strong solution of (B.1).

From the obtained convergence result, a confidence interval can be estimated for the numerical solution. In practice, this gives an estimate of how accurate the numerical solution is for a given simulation and can, furthermore, be used as a guide to determine the required size of the discretization of the Euler-Maruyama method.

The discontinuous element in the drift coefficient comes from the *sgn*-function. The sgn-function is interesting since it often appears in models of systems with sliding mode control [15, 16]. In these models, the discontinuity of the coefficients is often not bounded, which is required in almost all the results provided so far for SDEs with discontinuous coefficients.

In the following, we present the one-dimensional SDE.

1.1 Stochastic Differential Equation

A general one dimensional SDE is given by

$$dx_t = b(t, x_t)dt + \sigma(t, x_t)dW_t, \quad x_0 = c_0 \tag{B.2}$$

where $x = x_t : [0, T] \to \mathbb{R}$ is an $\mathbb{R}$-valued stochastic process, $b, \sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ are the drift and diffusion coefficient of x, $W = W_t$ is an $\mathbb{R}$-valued Wiener process, and c_0 a given constant.

This paper focuses on the special case of (B.2), where $\sigma(t, x_t) = 1$ and $b(t, x_t) = -k\,\mathrm{sgn}(x_t)$ with $k > 0$ a control gain and the sgn-function defined by

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{if} \quad x < 0 \\ 0 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x > 0 \end{cases} . \tag{B.3}$$

Thus, we consider the SDE

$$dx_t = -k\,\mathrm{sgn}(x_t)dt + dW_t, \quad x_0 = c_0 \tag{B.4}$$

with c_0 given.

In the sequel, the Euler-Maruyama method is presented and applied to (B.4).

2 The Euler-Maruyama Method

The Euler-Maruyama method is used to approximate solutions to SDEs of the type given in (B.2), by discretizing the time interval $[0, T]$ in steps $0 < t_1 < \cdots < t_n < t_{n+1} \cdots < t_N$ with $N = \left\lceil \frac{T}{h} \right\rceil$, where $h = t_{n+1} - t_n$ is a fixed time-step. For each step $t_n = nh$ a recursion gives an approximation to the solution to (B.2) by the following method,

$$x_{n+1} = x_n + hb(t_n, x_n) + \sigma(t_n, x_n)\Delta W_n \tag{B.5}$$

where $x_{t_i} = x_i$ and $\Delta W_n = W(t_{n+1}) - W(t_n)$ is i.i.d. normal with mean zero and variance h, which we denote by $\Delta W_n \sim N(0,h)$. Thus, given an initial condition $x_0 = c_0$, it is possible from (B.5) to approximate a solution to (B.2) with the sequence $\{x_n\} = \{x_{t_n} : n \in \{1,2,\ldots,N\}\}$, which, if the drift and diffusion coefficient are measurable, satisfy a Lipschitz condition and a growth bound, converges strongly to the solution of (B.2) as $h \to 0$, [7, Theorem 9.6.2]. Hence, for SDEs with discontinuous drift we cannot expect, in general, that the Euler-Maruyama method produces meaningful results, but as mentioned previously, under certain relaxed conditions on the drift and diffusion coefficients weak convergence and convergence in probability are proven.

In the following section, convergence aspects are investigated when $\{t_n\}$ are dyadic points.

3 Convergence in Dyadic Points

The Euler-Maruyama method is applied to (B.4) with dyadic points as evolution points. It is proven that the simulated solutions produced by the Euler-Maruyama method converge uniformly almost surely in each dyadic point for decreasing length of the time-step. That is, in the following are considered a sequence of simulations, produced with decreasing length of the time-step for a realization of the Wiener process. First, the notation is stated.

The Euler-Maruyama method is applied to (B.4) in the time interval $[0,T]$ with $T = 1$ and time-steps $h_n = 2^{-n}$. For a fixed $n \in \mathbb{N}$, that is a fixed h_n, the method produces steps $x^n(t_j^n) = x_j^n$ for times $0 < t_1^n < \cdots < t_j^n < \cdots < t_{2^n}^n = 1$, where $t_j^n = j \cdot h_n$. Therefore, the sequence $\{x_j^n\}_j = \{x^n(t_j^n) : j \in \{1,2,\ldots,2^n\}\}$ is generated by the recursion

$$x_{j+1}^n = x_j^n - h_n k \operatorname{sgn}(x_j^n) + \Delta W_j^n \,, \tag{B.6}$$

with time-step h_n, an initial condition $x_0^n = c_0$ and $\Delta W_j^n = W_{j+1}^n - W_j^n$, where $W_j^n = W(t_j^n)$ is a realization point from the Wiener process. It is assumed that the initial condition c_0 is a fixed constant.

The definition of h_n gives that a specific dyadic point $t_j^n = jh_n \in [0,T]$ is represented in all sequences $\{t_j^m\}_j = \{t_j^m : j \in \{1,2,\ldots,2^m\}\}$ for $m > n$ since $t_j^n = jh_n = 2jh_{n+1} = 2^{m-n}jh_m = t_{2^{m-n}j}^m$. Figure B.1 shows an example of elements from the sequence $\{x_j^n\}_j$ and $\{x_j^{n+1}\}_j$. Note in particular that x_{j+1}^n is produced from x_j^n, while $x_{2(j+1)}^{n+1}$ is produced from x_{2j+1}^{n+1}, so even though the Euler-Maruyama method uses the same realization of the Wiener process and the same inital condition, the simulated solutions are different.

To investigate convergence in dyadic points, the sequence

$$\{x_{2^p j}^{n+p}\}_p = \{x^{n+p}(2^p jh_{n+p}) : p \in \{0,1,\ldots\}\} \tag{B.7}$$

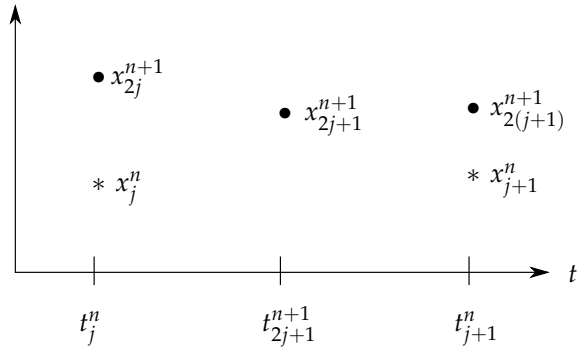

Figure B.1: Elements from the sequences $\{x^n\}_j$ and $\{x^{n+1}\}_j$.

is considered with an initially fixed $n \in \mathbb{N}$ and a fixed dyadic point $t_j^n = jh_n$, that is, realization points $x^{n+p}(t_j^n)$ for $p \geq 0$, produced at the same dyadic point t_j^n but from simulations with decreasing length of the time-step. We will show that (B.7) is a Cauchy sequence with the probability tending to one as $n \to \infty$.

The first two elements in the sequence $\{x_{2^p(j+1)}^{n+p}\}_p$ are

$$\begin{aligned} x_{j+1}^n &= x_j^n - h_n k \operatorname{sgn}(x_j^n) + W_{j+1}^n - W_j^n \\ &= x_j^n - 2h_{n+1} k \operatorname{sgn}(x_j^n) + W_{2(j+1)}^{n+1} - W_{2j}^{n+1} \end{aligned} \tag{B.8}$$

and

$$\begin{aligned} x_{2(j+1)}^{n+1} &= x_{2j+1}^{n+1} - h_{n+1} k \operatorname{sgn}(x_{2j+1}^{n+1}) + W_{2(j+1)}^{n+1} - W_{2j+1}^{n+1} \\ &= x_{2j}^{n+1} - h_{n+1} k \operatorname{sgn}(x_{2j}^{n+1}) + W_{2j+1}^{n+1} - W_{2j}^{n+1} \\ &\quad - h_{n+1} k \operatorname{sgn}(x_{2j+1}^{n+1}) + W_{2(j+1)}^{n+1} - W_{2j+1}^{n+1} \\ &= x_{2j}^{n+1} - h_{n+1} k \left(\operatorname{sgn}(x_{2j}^{n+1}) + \operatorname{sgn}(x_{2j+1}^{n+1}) \right) + W_{2(j+1)}^{n+1} - W_{2j}^{n+1} \, . \end{aligned}$$

The equality in (B.8) follows since $h_n = 2^p h_{n+p}$ and $W(jh_n) = W_j^n = W_{2^p j}^{n+p}$.

Hence, the difference between the two successive elements in the sequence $\{x_{2^p(j+1)}^{n+p}\}_p$ is

$$\begin{aligned} x_{2(j+1)}^{n+1} - x_{j+1}^n &= x_{2j}^{n+1} - x_j^n - h_{n+1} k \Big(\operatorname{sgn}(x_{2j}^{n+1}) + \operatorname{sgn}(x_{2j+1}^{n+1}) - 2 \operatorname{sgn}(x_j^n) \Big) \\ &= x_{2j}^{n+1} - x_j^n - h_{n+1} k Q_j^n \end{aligned} \tag{B.9}$$

where $Q_j^n = \operatorname{sgn}(x_{2j}^{n+1}) + \operatorname{sgn}(x_{2j+1}^{n+1}) - 2\operatorname{sgn}(x_j^n)$. Depending on the values of the sign-functions, different situations appear. In the sequel, this is investigated. For abbreviation, $\operatorname{sgn}(x_b^a)$ is denoted by s_b^a.

1. If $s_{2j}^{n+1}, s_{2j+1}^{n+1}, s_j^n$ share the same sign then

$$x_{2(j+1)}^{n+1} - x_{j+1}^n = x_{2j}^{n+1} - x_j^n .$$

2. If $s_{2j}^{n+1} = s_{2j+1}^{n+1} = -s_j^n \neq 0$ then

$$x_{2(j+1)}^{n+1} - x_{j+1}^n = x_{2j}^{n+1} - x_j^n - 4h_{n+1}k\,\mathrm{sgn}(Q_j^n)$$

where $\mathrm{sgn}(x_{2j}^{n+1} - x_j^n) = \mathrm{sgn}(Q_j^n)$. Therefore,

$$|x_{2(j+1)}^{n+1} - x_{j+1}^n| \leq |x_{2j}^{n+1} - x_j^n| . \tag{B.10}$$

3. If $s_{2j+1}^{n+1} = s_j^n = -s_{2j}^{n+1} \neq 0$ then

$$x_{2(j+1)}^{n+1} - x_{j+1}^n = x_{2j}^{n+1} - x_j^n - 2h_{n+1}k\,\mathrm{sgn}(Q_j^n) ,$$

and

$$\begin{aligned} (x_{2j}^{n+1} - x_j^n) &\leq 0 \text{ for } \mathrm{sgn}(Q_j^n) = -1 , \\ (x_{2j}^{n+1} - x_j^n) &\geq 0 \text{ for } \mathrm{sgn}(Q_j^n) = 1 . \end{aligned}$$

Both of the above options imply (B.10).

4. If $s_{2j}^{n+1} = s_j^n = -s_{2j+1}^{n+1} \neq 0$ then

$$x_{2(j+1)}^{n+1} - x_{j+1}^n = x_{2j}^{n+1} - x_j^n - 2h_{n+1}k\,\mathrm{sgn}(Q_j^n) .$$

In the following lemma, case 4) is studied in detail.

Lemma 3.1. *For case 4): If*

$$|x_{2j}^{n+1} - x_j^n| \geq c\sqrt[4]{h_n} \tag{B.11}$$

where $c > k$ is a fixed constant, then either

$$|x_{2(j+1)}^{n+1} - x_{j+1}^n| \leq |x_{2j}^{n+1} - x_j^n| \tag{B.12}$$

or

$$|\Delta W_{2j}^{n+1}| > \frac{k}{2}(\sqrt[4]{h_n} - h_n) . \tag{B.13}$$

Proof. If

$$\begin{aligned} x_{2j}^{n+1} - x_j^n &\leq 0 \text{ for } s_j^n = 1 \text{ or} \\ x_{2j}^{n+1} - x_j^n &\geq 0 \text{ for } s_j^n = -1 \end{aligned}$$

then (B.12) is valid. In the following, the converse cases are considered.

If

$$x_{2j}^{n+1} - x_j^n > 0 \text{ where } s_{2j}^{n+1} = s_j^n = 1 \,,$$

then by (B.11)

$$x_{2j}^{n+1} - x_{2j+1}^{n+1} > c\sqrt[4]{h_n} \,,$$

and similar if

$$x_{2j}^{n+1} - x_j^n < 0 \text{ where } s_{2j}^{n+1} = s_j^n = -1$$

then

$$x_{2j+1}^{n+1} - x_{2j}^{n+1} > c\sqrt[4]{h_n} \,.$$

Both situations imply that

$$|x_{2j}^{n+1} - x_{2j+1}^{n+1}| > c\sqrt[4]{h_n} \,.$$

From (B.6)

$$|x_{2j}^{n+1} - x_{2j+1}^{n+1}| \leq |h_{n+1}k| + |\Delta W_{2j}^{n+1}| \,,$$

so

$$\begin{aligned} |\Delta W_{2j}^{n+1}| &\geq |x_{2j}^{n+1} - x_{2j+1}^{n+1}| - |h_{n+1}k| \\ &> c\sqrt[4]{h_n} - h_{n+1}k = c\sqrt[4]{h_n} - h_n\frac{k}{2} \\ &> \frac{k}{2}\left(\sqrt[4]{h_n} - h_n\right) \,. \end{aligned}$$

Hence, for case 4), (B.11) imply either (B.12) or (B.13). □

In the following lemma, it is proven that the probability that (B.13) is the case for more than finitely many p, converges to zero as $n \to \infty$.

Lemma 3.2. *The probability*

$$P\Big(\exists p \geq 0 : |\Delta W_{2j}^{n+p}| > \frac{k}{2}(\sqrt[4]{h_{n+p-1}} - h_{n+p-1})\Big) \to 0 \text{ as } n \to \infty \,.$$

Proof. The properties of the Wiener process (i.e. $\Delta W_j^n \sim N(0, h_n)$) imply that

$$\begin{aligned} P\Big(|\Delta W_{2j}^{n+p}| > \frac{k}{2}(\sqrt[4]{h_{n+p-1}} - h_{n+p-1})\Big) < & \quad P\left(|\Delta W_{2j}^{n+p}| > \frac{\frac{k}{2}\sqrt[4]{h_{n+p-1}}}{2}\right) \\ < & \quad A\exp\left(-B\frac{1}{\sqrt[4]{h_{n+p-1}}}\right) \end{aligned} \tag{B.14}$$

where A and B are positive constants and B is proportional to the control gain k, i.e. $B = \frac{k}{\sqrt{2}}$. Due to the independent increments property of the Wiener process

$$\begin{aligned} P\Big(|\Delta W_{2j}^{n+p}| \leq \frac{k}{2}(\sqrt[4]{h_{n+p-1}} - h_{n+p-1})\ \forall j \in \{0,1,\ldots,2^n-1\}\Big) & \\ \geq & \quad \left(1 - A\exp\left(-B\frac{1}{\sqrt[4]{h_{n+p-1}}}\right)\right)^{2^n} \\ = & \quad \left(1 - A\exp\left(-B\sqrt[4]{2}^{n+p-1}\right)\right)^{2^n}. \end{aligned}$$

We use the fact that for large n

$$\begin{aligned} \log\left(\left(1 - A\exp\left(-B\sqrt[4]{2}^{n+p-1}\right)\right)^{2^n}\right) = & \quad 2^n \log\left(1 - A\exp\left(-B\sqrt[4]{2}^{n+p-1}\right)\right) \\ \approx & \quad -2^n A\exp\left(-B\sqrt[4]{2}^{n+p-1}\right), \end{aligned}$$

so

$$\begin{aligned} P\Big(|\Delta W_{2j}^{n+p}| \leq \frac{k}{2}(\sqrt[4]{h_n} - h_n)\forall j \in \{0,1,\ldots,2^n-1\}\Big) & \\ \geq & \quad \exp\left(-2^n A\exp\left(-B\sqrt[4]{2}^{n+p-1}\right)\right) \\ = & \quad \exp\left(-A\exp\left(-B\sqrt[4]{2}^{n+p-1} + n\log(2)\right)\right) \\ \geq & \quad \exp\left(-A\exp\left(-\frac{B}{2}\sqrt[4]{2}^{n+p-1}\right)\right). \end{aligned}$$

Thus, the probability that there exists a $p \geq 1$ such that

$$|\Delta W_{2j}^{n+p}| > \frac{k}{2}(\sqrt[4]{h_{n+p-1}} - h_{n+p-1})$$

is bounded from above by

$$\begin{aligned}\sum_{p=1}^{\infty}\left(1-\exp\left(-A\exp\left(-\frac{B}{2}\sqrt[4]{2}^{n+p-1}\right)\right)\right) &\approx \sum_{p=1}^{\infty} A\exp\left(-\frac{B}{2}\sqrt[4]{2}^{n+p-1}\right)\\ &\leq \sum_{p=0}^{\infty} A\exp\left(-\frac{B}{2}(n+p)\right)\\ &= A\exp\left(-\frac{nB}{2}\right)\sum_{p=0}^{\infty}\exp\left(-\frac{B}{2}\right)^{p}\\ &= \frac{A\exp\left(-n\frac{B}{2}\right)}{1-\exp\left(-\frac{B}{2}\right)} \quad := \quad P_n\,,\end{aligned} \tag{B.15}$$

which converges to zero for increasing n. □

Lemma 3.3. *For fixed n, if*

$$|\Delta W_{2j}^{n+1}| \quad \leq \quad \frac{k}{2}(\sqrt[4]{h_n}-h_n)\,, \tag{B.16}$$

for all $j \in \{0,1,\ldots,2^n-1\}$, *then*

$$|x_{2j}^{n+1}-x_j^n| \quad < \quad \hat{c}\sqrt[4]{h_n}\,, \tag{B.17}$$

for all $j \in \{0,1,\ldots,2^n-1\}$, *where* $\hat{c}$ *is a constant.*

Proof. The initial condition is the same for all simulations so initially the distance $|x_0^{n+1}-x_0^n|$ is equal to zero. Equation (B.9) shows that the increment that happens from simulation in one dyadic point to simulation in the next dyadic points, is bounded by $4h_{n+1}k = 2h_nk$. Hence, it might happen that at some dyadic point j

$$|x_{2j}^{n+1}-x_j^n| \geq c\sqrt[4]{h_n}\,. \tag{B.18}$$

However, if (B.16) is valid, lemma 3.1 and case 1) - 3) imply that the distance $|x_{2(j+1)}^{n+1}-x_{j+1}^n|$ is smaller than or equal to the previous one, i.e.

$$|x_{2(j+1)}^{n+1}-x_{j+1}^n| \leq |x_{2j}^{n+1}-x_j^n|\,.$$

That is, there exists a constant $\hat{c}$ such that for all dyadic points, (B.17) is valid. Figure B.2 illustrates an example of the evolution in the distance $|x_{2j}^{n+1}-x_j^n|$ as the simulation in dyadic points goes on. □

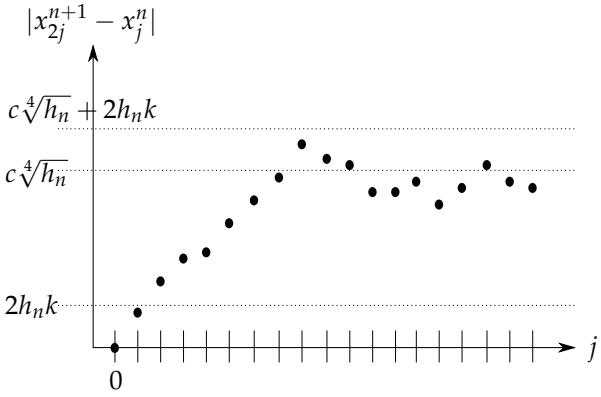

Figure B.2: Evolution of distance between two successive elements in the sequence $\{x_{2^p j}^{n+p}\}_p$ for increasing dyadic points.

From lemma 3.2 and 3.3 the following lemma can be proven.

Lemma 3.4. *The sequence $\{x_{2^p j}^{n+p}\}_p$ is a Cauchy sequence with probability $1 - P_n$, with P_n given in* (B.15)*.*

Proof. For fixed n assume without loss of generality that $p_1 > p_2$ and consider

$$\begin{aligned} |x_{2^{p_1} j}^{n+p_1} - x_{2^{p_2} j}^{n+p_2}| &\leq \sum_{i=n+p_2+1}^{n+p_1} |x_{2^{i-n} j}^{i} - x_{2^{i-n-1} j}^{i-1}| \\ &\leq \hat{c} \sum_{i=n+p_2+1}^{n+p_1} \sqrt[4]{h_{i-1}}\,. \end{aligned} \tag{B.19}$$

From lemma 3.2 and 3.3 inequality (B.19) holds with probability $1 - P_n$. Furthermore,

$$\begin{aligned} \hat{c} \sum_{i=n+p_2+1}^{n+p_1} \sqrt[4]{h_{i-1}} &\leq \hat{c} \sqrt[4]{2}^{-(n+p_2)} \sum_{i=0}^{\infty} \sqrt[4]{2}^{-i} \\ &= \hat{c} \frac{\sqrt[4]{2}^{-(n+p_2-1)}}{\sqrt[4]{2} - 1}\,. \end{aligned} \tag{B.20}$$

The value in (B.20) obtains its maximum for $p_2 = 0$. Since $p_2 \geq 0$, the tail sum converges to zero as $n \to \infty$ no matter the value of p_2. Thus, the sequence $\{x_{2^p j}^{n+p}\}_p$ is Cauchy with probability $1 - P_n$. □

Following theorem proves that lemma 3.4 implies almost surely convergence of the Euler-Maruyama simulations.

Theorem 3.5. *The sequence $\{x_{2^p j}^{p}\}_p$ converges almost surely.*

Proof. The Cauchy property implies that if $\{x_{2^{p_j}}^{n+p}\}_p$ is Cauchy then $\{x_{2^{p_j}}^{m+p}\}_p$ is Cauchy for all $n \geq m$. Therefore, since $\{x_{2^{p_j}}^{n+p}\}_p$ is Cauchy, the sequence $\{x_{2^{p_j}}^{p}\}_p$ is Cauchy. Thus,

$$P(\{x_{2^{p_j}}^{n+p}\}_p \text{ is Cauchy }) \quad \leq \quad P(\{x_{2^{p_j}}^{p}\}_p \text{ is Cauchy })$$

for all $n \geq 0$. However, from lemma 3.2 and lemma 3.4

$$P(\{x_{2^{p_j}}^{n+p}\}_p \text{ is Cauchy }) \to 1 \text{ for } n \to \infty$$

and therefore,

$$P(\{x_{2^{p_j}}^{p}\}_p \text{ is Cauchy }) = 1 \, .$$

Hence, since $\{x_{2^{p_j}}^{p}\}_p$ is Cauchy with probability one, it follows that $\{x_{2^{p_j}}^{p}\}_p$ converges almost surely. □

In the following section, we will discuss the consequences of this.

4 Rate of Convergence

As mentioned in the introduction, [14] proves a weak rate of convergence of the Euler-Maruyama method for an example similar to (B.4). Following, the strong rate of convergence for (B.4) is presented. Given a numerical solution and the corresponding probability P_n, a confidence interval can be deduced. For the numerical solution, the probability that its distance to the limit solution in each dyadic point is above

$$\hat{c}\frac{\sqrt[4]{2}^{-(n-1)}}{\sqrt[4]{2}-1} \, , \tag{B.21}$$

is less than P_n. Thus, this represents a $(1-2P_n)\cdot 100\%$ confidence interval. Furthermore, with $\{x\}$ denoting the "true" solution

$$P\left(|x(T)-x^n(T)| \geq \hat{c}\frac{\sqrt[4]{2}^{-(n-1)}}{\sqrt[4]{2}-1}\right) \leq P_n \, ,$$

and since by construction invariantly $|x(T)-x^n(T)| \leq 2k$ we have

$$\mathbb{E}[|x(T)-x^n(T)|] \quad \leq \quad \hat{c}\frac{\sqrt[4]{2}^{-(n-1)}}{\sqrt[4]{2}-1} + 2kP_n \, .$$

Altogether, with T equal to a dyadic point,

$$\begin{aligned}\mathbb{E}[|x(T)-x^n(T)|] \quad &\leq \quad \hat{c}\frac{\sqrt[4]{2}^{-(n-1)}}{\sqrt[4]{2}-1} + 2kP_n \\ &\leq \quad C\left(h_n^{\frac{1}{4}} + h_n^{\frac{B}{2\log(2)}}\right) \, ,\end{aligned}$$

where $C = \max\{\frac{\hat{c}\sqrt[4]{2}}{\sqrt[4]{2}-1}, \frac{2kA}{1-\exp(-\frac{B}{2})}\} > 0$. Therefore, it follows that the Euler-Maruyama method applied to (B.4) converges strongly with order $\gamma = \min\{\frac{1}{4}, \frac{B}{2\log(2)}\}$. Since $\frac{B}{2\log(2)} \approx \frac{k}{2}$, the convergence order is at least $\frac{1}{4}$ for values of $k > \frac{1}{2}$.

Corollary 4.1. *The sequence $\{x^p_{2^p j}\}$ converges to a strongly unique solution.*

Proof. From theorem 3.5, the sequence $\{x^p_{2^p j}\}$ converges in the space of real numbers, which is complete, so there exists a limit random variable x to which $\{x^p_{2^p j}\}$ converges as $p \to \infty$, i.e.

$$\{x^p_{2^p j}\} \to x \text{ as } p \to \infty \tag{B.22}$$

with probability one. It is known from [11] that a limit value y exists, such that

$$\{x^p_{2^p j}\} \to y \text{ as } p \to \infty \tag{B.23}$$

in probability. From [5], the solution to (B.4) is strongly unique and hence, (B.22) and (B.23) imply that $x = y$ with probability one. □

In the following section, a numerical example is presented together with the corresponding confidence interval for the simulated solutions.

5 Numerical Example

The Euler-Maruyama method is applied to the SDE given in (B.4) with initial condition $c_0 = 0$ and with the control gain $k = 1$. The simulations are repeated for increasing values of n up to $n = 22$. A selection of simulated values of $x^n(t)$ is given in Table B.1. The results indicate that the simulated solutions tend to a limit solution in the dyadic points, since the difference between two simulated solutions in general decrease as n increase. This confirms the theoretical result presented in lemma 3.3, i.e. the difference in dyadic points between two simulations is bounded from above by a constant dependent on the time-step.

For $A = \exp(2)$, $B = \frac{k}{\sqrt{2}}$ and $\hat{c} = 1$, a 95% confidence interval is obtained for $n = 20$, given by

$$[x^{20}(t) - 0.1964, x^{20}(t) + 0.1964] .$$

The estimate in (B.14) holds for the given value of A and B. Furthermore, it is reasonable to choose $\hat{c} = 1$ for $k = 1$, since the required difference between $\hat{c}$ and k decreases for increasing n.

In Figure B.3, $x^{20}(t)$ is plotted together with the upper- and lower bound for the confidence interval. Furthermore, the simulation $x^{22}(t)$ is plotted over $x^{20}(t)$. The zoomed area shows that $x^{22}(t)$ indeed stays inside the confidence interval as is expected for values $n \geq 20$.

$x^n(t) \setminus t$	0.25	0.50	0.75	1.00
$x^4(t)$	−1.1254	−0.5891	0.5705	0.9031
$x^5(t)$	−1.1567	−0.6204	0.4143	0.7469
$x^6(t)$	−1.1410	−0.6047	0.4299	0.7625
$x^7(t)$	−1.1645	−0.6282	0.4377	0.7703
$x^8(t)$	−1.1606	−0.6243	0.4338	0.7664
$x^9(t)$	−1.1547	−0.6184	0.4201	0.7527
$x^{10}(t)$	−1.1596	−0.6233	0.4192	0.7518
$x^{11}(t)$	−1.1591	−0.6228	0.4187	0.7513
$x^{12}(t)$	−1.1589	−0.6226	0.4170	0.7496
$x^{13}(t)$	−1.1595	−0.6232	0.4171	0.7497
$x^{14}(t)$	−1.1597	−0.6234	0.4169	0.7495
$x^{17}(t)$	−1.1595	−0.6232	0.4174	0.7500
$x^{20}(t)$	−1.1595	−0.6232	0.4174	0.7501
$x^{22}(t)$	−1.1595	−0.6232	0.4174	0.7501

Table B.1: Simulation results in four dyadic points for increasing values of n.

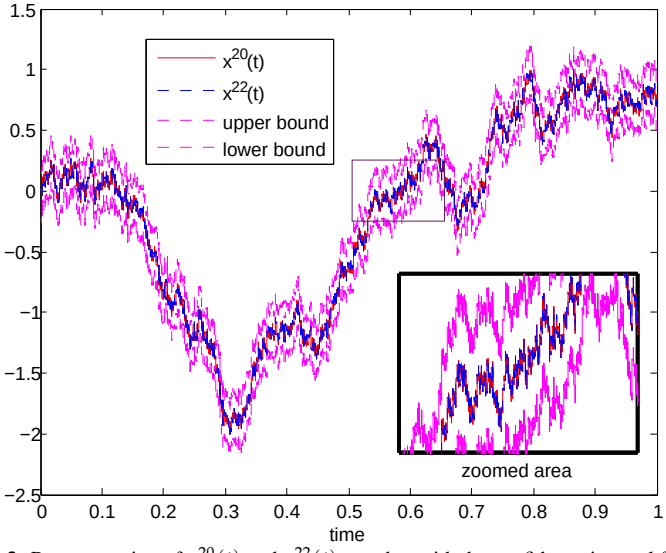

Figure B.3: Representation of $x^{20}(t)$ and $x^{22}(t)$ together with the confidence interval for $x^{20}(t)$.

6 Conclusion

This paper focuses on a simple SDE from a class of SDEs with discontinuous drift coefficients. We provide a strong rate of convergence in probability of the Euler-Maruyama method for this particular example and we prove that numerical solutions produced by the Euler-Maruyama method converge almost surely to a unique strong solution of the SDE. A numerical example illustrates that the Euler-Maruyama simulated solutions indeed converge as n increase.

Future work is to consider SDEs with discontinuous unbounded drift coefficients. For example, one could implement the drift $b(t, x_t) = -x_t - \text{sgn}(x_t)$ in an SDE and investigate convergence aspects of the Euler-Maruyama method. Another step is to generalize the result to multidimensional SDEs with discontinuous drift and prove convergence for these cases.

References

[1] K. Ito, "On stochastic differential equations," *Mem. Amer. Math. Soc.*, vol. 1951, no. 4, p. 51, 1951.

[2] R. L. Stratonovich, "A new representation for stochastic integrals and equations," *SIAM J. Control*, vol. 4, pp. 362–371, 1966.

[3] A. S. Cherny and H.-J. Engelbert, *Singular stochastic differential equations*, ser. Lecture Notes in Mathematics. Berlin: Springer-Verlag, 2005, vol. 1858.

[4] A. K. Zvonkin, "A transformation of the phase space of a diffusion process that will remove the drift," *Mat. Sb. (N.S.)*, vol. 93(135), pp. 129–149, 152, 1974.

[5] A. J. Veretennikov, "On strong solutions and explicit formulas for solutions of stochastic integral equations," *Math. USSR Sb.*, vol. 39, no. 3, pp. 387–403, 1981. [Online]. Available: http://dx.doi.org/10.1070/SM1981v039n03ABEH001522

[6] G. Maruyama, "Continuous Markov processes and stochastic equations," *Rend. Circ. Mat. Palermo (2)*, vol. 4, pp. 48–90, 1955.

[7] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, ser. Applications of Mathematics (New York). Berlin: Springer-Verlag, 1992, vol. 23.

[8] C. Yuan and X. Mao, "Convergence of the Euler-Maruyama method for stochastic differential equations with Markovian switching," *Math. Comput. Simulation*, vol. 64, no. 2, pp. 223–235, 2004. [Online]. Available: http://dx.doi.org/10.1016/j.matcom.2003.09.001

[9] Y. Wang and C. Yuan, "Convergence of the Euler-Maruyama method for stochastic differential equations with respect to semimartingales," *Appl. Math. Sci. (Ruse)*, vol. 1, no. 41-44, pp. 2063–2077, 2007.

[10] H. Li, L. Xiao, and J. Ye, "Strong predictor-corrector Euler-Maruyama methods for stochastic differential equations with Markovian switching," *J. Comput. Appl. Math.*, vol. 237, no. 1, pp. 5–17, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.cam.2012.07.001

[11] I. Gyöngy and N. Krylov, "Existence of strong solutions for Itô's stochastic equations via approximations," *Probab. Theory Related Fields*, vol. 105, no. 2, pp. 143–158, 1996. [Online]. Available: http://dx.doi.org/10.1007/BF01203833

[12] K. S. Chan and O. Stramer, "Weak consistency of the Euler method for numerically solving stochastic differential equations with discontinuous coefficients," *Stochastic Process. Appl.*, vol. 76, no. 1, pp. 33–44, 1998. [Online]. Available: http://dx.doi.org/10.1016/S0304-4149(98)00020-9

[13] L. Yan, "The Euler scheme with irregular coefficients," *Ann. Probab.*, vol. 30, no. 3, pp. 1172–1194, 2002. [Online]. Available: http://dx.doi.org/10.1214/aop/1029867124

[14] A. Kohatsu-Higa, A. Lejay, and K. Yasuda, "On Weak Approximation of Stochastic Differential Equations with Discontinuous Drift Coefficient." [Online]. Available: http://hal.inria.fr/hal-00670123

[15] V. I. Utkin, J. Guldner, and J. Shi, *Sliding mode control in electromechanical systems*, ser. The Taylor & Francis systems and control book series. London, Philadelphia, PA: Taylor & Francis, 1999.

[16] F. H. Clarke and R. B. Vinter, "Stability analysis of sliding-mode feedback control," *Control Cybernet.*, vol. 38, no. 4A, pp. 1169–1192, 2009.

References

Paper C

Investigation of Random Switching Driven by a Poisson Point Process

Maria Simonsen, Henrik Schiøler and John Leth

The paper has been published in the
proceedings of IEEE Conference on Control Applications, 2015.

The layout has been revised, and small editorial changes have been made. Content relevant changes, if any, are marked with explicit footnotes.

Abstract

This paper investigates the switching mechanism of a two-dimensional switched system when the switching events are generated by a Poisson point process. A model, in the shape of a stochastic process, for such a system is derived and the distribution of the trajectory's position is developed together with marginal density functions for the coordinate functions. Furthermore, the joint probability distribution is given explicitly.

1 Introduction

The behaviour of dynamical systems is suitably modelled mathematically with differential equations, either ordinary (ODE), partial (PDE) or stochastic (SDE). Examples are found in thermodynamics, mechanics, chemistry, economics and electronics, [1]. The control of dynamic systems alters the *open loop* differential equations into *closed loop dynamics*, which in some cases, when switching is applied, introduces discontinuities to the mathematical model. When discontinuities are constrained to some surface, i.e. a subset of the state space of lower dimension, we call this subset a *switching surface*. Altogether, the constellation of system and controller falls under the category of *switched system*, [2]. The purpose of switching control, in some cases, is to confine the system state to the switching surface and in turn to obtain so called *sliding mode* behaviour/dynamics with desired properties, [3]. Advantages of switching control as compared to continuous/linear controllers are found in: robustness to disturbances, simplicity of actuator design and, in some cases, savings in actuator energy consumption. Recent applications of sliding mode control are found in [4], [5] and [6] indicating mechanics/robotics to be a distinctively active application area for this methodology.

Systems with state dependent switching have been modelled with different approaches in the existing literature, where the lack of continuity imposes a significant difficulty as the problem is rendered outside the scope of the classical *initial value* setting. One early and general approach is suggested in [7], introducing the *Filippov solution* as the solution to a convex version of a differential inclusion corresponding to the vector fields with discontinuous dynamics. The sliding velocity of the Filippov solution on the sliding surface has been determined in [8] based on two regularizations. One of these regularizations implemented a small fixed delay into the deterministic system and deduced the sliding velocity under this restriction.

Probabilistic approaches utilize the smoothing effect of randomization, such as in [9] where the switching between two states is governed by a probability law. In that framework, the switching can occur spontaneously according to the probability law before the boundary of the local state space or at the boundary of the local state space

where the switching is enforced. The reverse framework is presented in [10] where systems, which might induce a time-delay in the switches are considered. Here the system flows for some time after crossing a local boundary before the switching takes place. In between these frameworks the setup in [11] can be found where switching noise is considered. In this case the switching is governed by an intensity function and the switching between subsystems occurs in a small area around the switching surface.

It may be discussed, whether randomization is in fact introduced to model inherent system randomness or simply to render the model equations mathematically tractable. It is argued in [11] that control system designers deliberatively introduce randomness to avoid undesired synchronization behaviour. Another example of inherent randomness is when random disturbances affect system behaviour. In this case, disturbance signals are often modelled in the domain of SDEs, [12], which in the switching case comprise *irregular drift*. As opposed to the ODE case, SDEs with irregular (discontinuous) drift are shown to comprise classical SDE solutions, i.e. existence and uniqueness of strong solutions are proven for bounded drift: [13], [14].

This paper presents a model of randomized switching driven by the Poisson point process, effectively acting as a switching delay. Contrary to the analysis in [8], the induced delay is stochastic and, accordingly, the solution is a stochastic process. We provide probabilistic results and compare the results with the Filippov solution provided for the deterministic ODE case. We narrow our focus and delimit ourselves to only a small region of the state space, which is explained in Figure C.1. The first subplot of Figure C.1 shows a classic sliding mode example where every trajectory is driven to the sliding surface and from there to the origin. The highlighted square on the left part of the plot illustrates the focus of this paper: the switching behaviour close to the switching surface away from the origin. For simplification, the sliding surface is rotated to a horizontal position, which is illustrated in the second subplot in Figure C.1. Additionally, investigations are done under the approximation that vector fields are constant above and below the switching surface, which is warranted by the continuity of vector fields on each side of the surface and the fact, that only a small portion of the state space is considered.

Even though the presented analysis is carried out in the two-dimensional framework, this setup can be extended to the n-dimensional setting to model systems with a switching surface of co-dimension one where the switching depends on the position according to this switching surface independently of the subsystem in $(n-1)$-dimensions.

For the two-dimensional framework, we deduce marginal density functions for the two coordinate functions, which are given in Theorem 3.1 and Corollary 3.2. Furthermore, a relation between the two coordinate functions is presented and an expression for the joint distribution function is given in Corollary 3.3. As indicated above, the

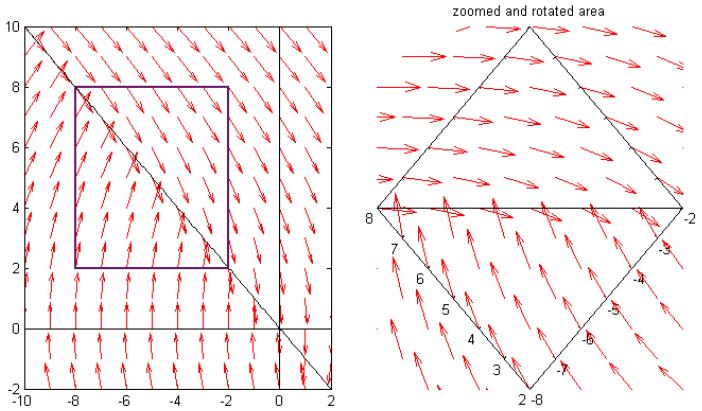

Figure C.1: The switching mechanism at the switching surface is extracted from a classic sliding mode example and rotated to the horizontal position.

practical relevance of this (initial) study is rooted in the attempt to describe (or more precisely to estimate) the collected behaviour of a family of physical systems connected via certain switching laws, which are governed by some degree of randomness. Furthermore, in a practical application setting the results of the paper can be used in a local analysis of such systems. In particular, the explicit description of the marginal density functions can be used to (locally) estimate the system behaviour.

2 A System with Constant Vector Fields

The framework for the switched system is a state space representation consisting of two constant vector fields. Let the state space $\mathbb{R}^2$, be partitioned in

$$\begin{aligned} S_0 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}, \\ S_1 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 < 0\}, \\ S_2 &= \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}. \end{aligned}$$

For $\mathbf{x} \in \mathbb{R}^2$ define $f_1(\mathbf{x}) = \mathbf{v}^- = (v_1^-, v_2^-)$ and $f_2(\mathbf{x}) = \mathbf{v}^+ = (v_1^+, v_2^+)$ where $\mathbf{v}^-$, $\mathbf{v}^+$ are constant vectors with the properties that $v_2^- > 0$ and $v_2^+ < 0$. Moreover, for technical convenience we assume that $\mathbf{v}^- \neq -\mathbf{v}^+$. Initially, consider the differential inclusion

$$\dot{\mathbf{x}} \in f(\mathbf{x}) \tag{C.1}$$

where

$$f(\mathbf{x}) = \left\{ \mathbf{v} \in \mathbb{R}^2 \mid \mathbf{v} = f_i(\mathbf{x}) \text{ if } \mathbf{x} \in \tilde{S}_i \text{ for } i = 1, 2 \right\} \tag{C.2}$$

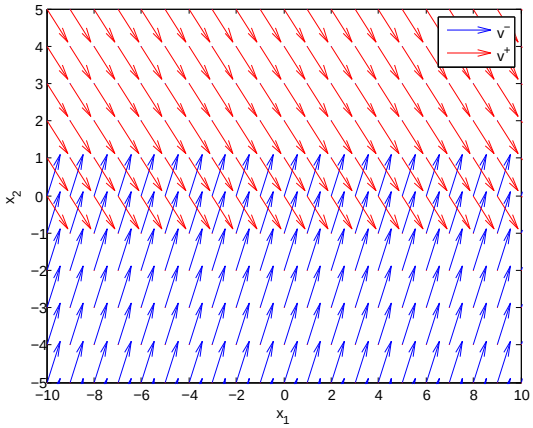

Figure C.2: The vector field given by the differential inclusion (C.1) with $\mathbf{v}^- = (1,2)$ and $\mathbf{v}^+ = (1,-1)$.

with $\tilde{S}_1 = S_1 \cup S_0$ and $\tilde{S}_2 = S_2 \cup S_0$. Figure C.2 gives an example of two constant vector fields with the required properties. Here $\mathbf{v}^- = (1,2)$ and $\mathbf{v}^+ = (1,-1)$. Given an initial condition, a solution to the differential inclusion (C.1) will evolve according to the vector field corresponding to the current position of the trajectory. When the trajectory reaches a position where $x_2 = 0$ it will stop, since no solution exists to (C.1) for $\mathbf{x} \in S_0$.

2.1 The Switching Surface and the Filippov Solution

Since (C.1) has no solution at the switching surface S_0, a convex version is considered which has a solution on S_0. Thus, consider the convex-hull

$$co(f(\mathbf{x})) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \notin S_0 \\ \{p\mathbf{v}^- + (1-p)\mathbf{v}^+ \ : \ p \in [0,1]\} & \text{if } \mathbf{x} \in S_0 \end{cases} .$$

A differential inclusion is then given by

$$\dot{\mathbf{x}}(t) \quad \in \quad co(f(\mathbf{x}(t))) \, . \tag{C.3}$$

Hence, $\mathbf{x}(t)$ is a Filippov solution to (C.1), if (C.3) holds almost everywhere. It follows that the unique Filippov solution at $\mathbf{x}(0) \in S_0$ must fulfil $\frac{d}{dt}x_2(t) = 0$. This implies that

$$p = \frac{v_2^+}{v_2^+ - v_2^-} \, , \tag{C.4}$$

and, therefore,

$$\frac{d}{dt}x_1(t) \quad = \quad pv_1^- + (1-p)v_1^+ = \frac{v_2^+ v_1^- - v_2^- v_1^+}{v_2^+ - v_2^-} \, . \tag{C.5}$$

The vector field provided by the Filippov solution is called the *Filippov velocity*.

In the sequel, we are aiming to obtain a relation between the two coordinate functions x_1 and x_2. We initiate this deterministically by a travelling wave approach.

2.2 The Travelling Wave

The constant vector fields imply that the trajectory moves with a constant velocity on the sliding surface. The velocity generated by the Filippov solution can be considered as the average velocity of the system. Let $\mathbf{C} = (C_1, 0)$ denote the Filippov velocity where C_1 is given by (C.5). We define local coordinates given by

$$(\eta_1, \eta_2) = \eta(\mathbf{x}, t) = \mathbf{x} - \mathbf{C}t \ .$$

The derivative of $\eta(\mathbf{x}, t)$ along $\mathbf{x}(t)$ is

$$\begin{aligned} \frac{d}{dt}\eta(\mathbf{x}(t), t) &= \frac{d}{dt}\mathbf{x}(t) - \mathbf{C} \ = \ co(f(\mathbf{x}(t))) - \mathbf{C} \\ &= \begin{cases} \left(\kappa v_2^-, v_2^-\right) & \text{for } \mathbf{x}(t) \in S_1 \\ (0,0) & \text{for } \mathbf{x}(t) \in S_0 \\ \left(\kappa v_2^+, v_2^+\right) & \text{for } \mathbf{x}(t) \in S_2 \end{cases} \end{aligned}$$

where p is given in (C.4) and $\kappa = \frac{v_1^- - v_1^+}{v_2^- - v_2^+}$. Recall that $\frac{d}{dt}x_2 = v_2^j$ for $j \in \{-, +\}$. Therefore,

$$\frac{dx_2}{d\eta_1} = \frac{\frac{d}{dt}x_2}{\frac{d}{dt}\eta_1} = \frac{v_2^j}{\kappa v_2^j} = \frac{1}{\kappa}$$

for $j \in \{-, +\}$. We remark that by taking the limits above, the equation is also valid for $\mathbf{x} \in S_0$. Hence, the sign of x_2 does not influence $\frac{dx_2}{d\eta_1}$ and the following relation between $\eta_1 = x_1 - C_1 t$ and x_2 is obtained,

$$x_2 \ = \ \frac{1}{\kappa}\eta_1 + c$$

where c is a constant. Under the assumption that the initial condition of the trajectory is $(x_1(t_0), x_2(t_0)) = (0, 0)$ with $t_0 = 0$, it follows that $c = 0$. Therefore,

$$x_2(t) \ = \ \frac{1}{\kappa}(x_1(t) - C_1 t) \ , \tag{C.6}$$

and a linear relation between $x_1(t)$ and $x_2(t)$ is given.

The relation in (C.6) is obtained in a deterministic setting. In the sequel we investigate the relation between the coordinate functions $x_1(t)$ and $x_2(t)$ when the switching mechanism is driven by a Poisson point process.

3 The Randomized Mechanism

In this section a stochastic process is constructed which realizations are concatenations of trajectories generated by the ordinary differential equations $\frac{d}{dt}\mathbf{x} = f_i(\mathbf{x})$, for $i = 1, 2$. The process is constructed such that the times where the trajectory switches between the two vector fields are defined by a Poisson point process. Thus, let $\{\tau_n\}$ denote a Poisson point process with intensity λ and let the stochastic process $\mathbf{x}(t)$ be defined piecewise as follows:

Model 1. *Initially, as the solution to*

$$\frac{d}{dt}\mathbf{x}(t) \quad = \quad f_1(\mathbf{x}(t)), \quad \mathbf{x}(t_0) = 0, \quad t_0 = 0$$

for $t \in [0, \tau_{m_1})$ *with* $\tau_{m_1} = \min\{\tau_n > 0\}$.

Thereafter, recursively for $i \in \{1, 2, \ldots\}$*, as the solution to*

$$\frac{d}{dt}\mathbf{x}(t) \quad = \quad f_{\chi(\tau_{m_i})}(\mathbf{x}(t)), \quad \mathbf{x}(\tau_{m_i}) = \lim_{t \nearrow \tau_{m_i}} \mathbf{x}(t)$$

for $t \in [\tau_{m_i}, \tau_{m_{i+1}})$ *with* $\tau_{m_{i+1}} = \min\{\tau_n > t_i\}$ *where* $t_i = \min\{t \in [\tau_{m_i}, \infty) | \mathbf{x}(t) = 0\}$ *and*

$$\chi(\tau_{m_i}) = \begin{cases} 1 & \text{if } x_2(\tau_{m_i}) < 0 \\ 2 & \text{if } x_2(\tau_{m_i}) > 0 \end{cases} .$$

We remark that the case $x_2(\cdot) = 0$ in the definition of $\chi(\cdot)$ is left unspecified since this case does not affect the stochastic process.

A trajectory generated by Model 1 has the property that it intersects the switching surface S_0 at t_i, that is, $x_2(t_i) = 0$ for all $i \in \{0, 1, 2, \ldots\}$ and $0 = t_0 < t_1 < t_2 < \cdots$. Furthermore, the i'th change in the discrete state occurs at $\tau_{m_i} = \min\{\tau_n > t_{i-1}\}$. Therefore, there might exist some $t \in \{\tau_n\}$ where no changing in states occur, i.e. in the cases where the trajectory has not reached to pass the switching surface, that is, $t_i - \tau_{m_i} > \tau_{m_i+1} - \tau_{m_i}$ for $\tau_{m_i}, \tau_{m_i+1} \in \{\tau_n\}$. Figure C.3 illustrates a version of the time evolution and the corresponding notation together with the discrete state values.

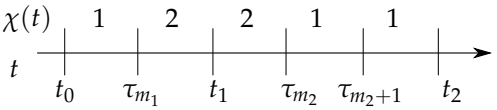

Figure C.3: The time evolution and the corresponding discrete state values. The behaviour of a trajectory resulting in these discrete states can be seen in Figure C.4.

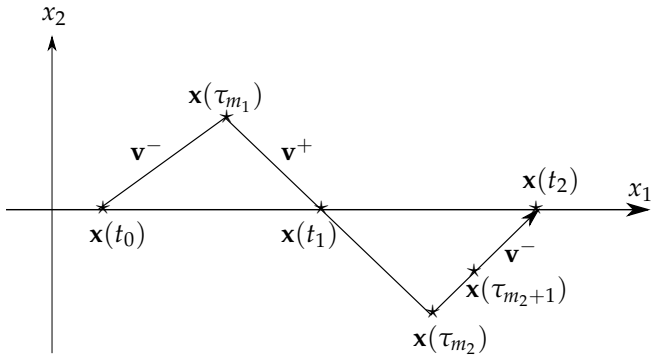

Figure C.4: The one-cycle behaviour of the trajectory for $t \in [t_0, t_2]$.

The interval $[t_0, t_2]$ comprises an entire cycle and for $t = t_2$ the process is renewed. This is illustrated in Figure C.4 where the one-cycle behaviour of the trajectory is shown.

The probabilistic behaviour of the trajectory based on its one-cycle behaviour is investigated in the following.

3.1 Probabilistic Investigations

The construction of the randomized mechanism and the corresponding one-cycle behaviour of the trajectory implies that the probabilistic behaviour of the system in the time intervals $[t_2, t_4], [t_4, t_6], \ldots$ is the same as in the time interval $[t_0, t_2]$, since the switching behaviour regenerates the initial position of the one-cycle, for every time t_{2i}, for $i = 1, 2, \ldots$. On this basis, probabilistic properties for the switching mechanism are developed.

Assume existence of a marginal density function for the coordinate function x_2 given by $\phi_{x_2}(x_2, t)$. Let $\mathbf{v}(t) = f(\mathbf{x}(t))$ where $f(\mathbf{x})$ is given by (C.2) and let $A \in \mathbb{R}$ be an open set. The positional probability for the coordinate function x_2 is given by

$$\begin{aligned} P(x_2(t) \in A, \mathbf{v}(t) = \mathbf{v}^j) &= \int_A Q^j(x_2, t)\phi_{x_2}(x_2, t)dx_2 \\ &= \int_A \Phi^j(x_2, t)dx_2 \end{aligned} \tag{C.7}$$

where $Q^j(x_2, t) = P(\mathbf{v}(t) = \mathbf{v}^j | x_2(t) = x_2)$ is the conditional probability of $\mathbf{v}(t)$ given $x_2(t) = x_2$ and $\Phi^j(x_2, t)$ is defined by the above expression.

Let h be a sufficiently small time-step and consider an approximation of the posi-

tional probability at time $t+h$,

$$P(x_2(t+h) \in A, \mathbf{v}(t+h) = \mathbf{v}^j) = (1-\lambda h) \int_{A-hv_2^j} \Phi^j(x_2,t)dx_2 + \lambda h \sum_i \int_{A-hv_2^i} P_{ij}(x_2)\Phi^i(x_2,t)dx_2 + o(h) \tag{C.8}$$

where $P_{ij}(x_2)$ is the probability of switching from the vector field $f_i(\mathbf{x})$ to the vector field $f_j(\mathbf{x})$, λ is the intensity of the Poisson point process, and $o(h)$ includes terms of h of degree two or higher.

Observe that $\Phi^j(x_2,t) - \Phi^j(x_2+hv_2^j,t) \approx -hv_2^j D_{x_2}[\Phi^j](x_2,t)$. By substituting this into (C.8), and in (C.9) and (C.10) below ignoring higher order terms of h, (C.8) can be expressed as

$$\begin{aligned} &P(x_2(t+h) \in A, \mathbf{v}(t+h) = \mathbf{v}^j) \\ \approx &(1-\lambda h) \int_{A-hv_2^j} \Phi^j(x_2,t) - \Phi^j(x_2+hv_2^j,t) + \Phi^j(x_2+hv_2^j,t)dx_2 \\ &+ \lambda h \sum_i \int_{A-hv_2^i} P_{ij}(x_2)\Big[\Phi^i(x_2,t) - \Phi^i(x_2+hv_2^i,t) + \Phi^i(x_2+hv_2^i,t)\Big]dx_2 \end{aligned} \tag{C.9}$$

$$\begin{aligned} \approx &(1-\lambda h) \int_{A-hv_2^j} -hv_2^j D_{x_2}[\Phi^j](x_2,t) + \Phi^j(x_2+hv_2^j,t)dx_2 \\ &+ \lambda h \sum_i \int_{A-hv_2^i} P_{ij}(x_2)\Big[-hv_2^i D_{x_2}[\Phi^i](x_2,t) + \Phi^i(x_2+hv_2^i,t)\Big]dx_2 \\ \approx &\int_{A-hv_2^j} -hv_2^j D_{x_2}[\Phi^j](x_2+hv_2^j,t) + (1-\lambda h)\Phi^j(x_2+hv_2^j,t)dx_2 \\ &+ \lambda h \sum_i \int_{A-hv_2^i} P_{ij}(x_2)\Phi^i(x_2+hv_2^i,t)dx_2\,. \end{aligned} \tag{C.10}$$

By changing variables,

$$\begin{aligned} P(x_2(t+h) \in A, \mathbf{v}(t+h) = \mathbf{v}^j) \\ \approx &\int_A -hv_2^j D_{x_2}[\Phi^j](x_2,t) + (1-\lambda h)\Phi^j(x_2,t)dx_2 \\ &+ \lambda h \sum_i \int_A P_{ij}(x_2-hv_2^i)\Phi^i(x_2,t)dx_2\,. \end{aligned}$$

From this

$$\begin{aligned}\int_A \Phi^j(x_2, t+h)dx_2 &= P(x_2(t+h) \in A, \mathbf{v}(t+h) = \mathbf{v}^j) \\ &\approx \int_A -hv_2^j D_{x_2}[\Phi^j](x_2,t) + (1-\lambda h)\Phi^j(x_2,t)dx_2 \\ &\quad + \lambda h \sum_i \int_A P_{ij}(x_2 - hv_2^i)\Phi^i(x_2,t)dx_2 \,.\end{aligned}$$

Rearranging of the terms gives

$$\begin{aligned}\int_A \Phi^j(x_2, t+h) - \Phi^j(x_2,t)dx_2 \approx \int_A &-hv_2^j D_{x_2}[\Phi^j](x_2,t) - \lambda h\Phi^j(x_2,t)dx_2 \\ &+ \lambda h \sum_i \int_A P_{ij}(x_2 - hv_2^i)\Phi^i(x_2,t)dx_2\end{aligned}$$

and, therefore, it follows that

$$\begin{aligned}D_t[\Phi^j](x_2,t) &= \lim_{h\to 0} \frac{\Phi^j(x_2,t+h) - \Phi^j(x_2,t)}{h} \\ &\approx -v_2^j D_{x_2}[\Phi^j](x_2,t) - \lambda\Phi^j(x_2,t) + \lambda \sum_i P_{ij}(x_2)\Phi^i(x_2,t) \,. \qquad \text{(C.11)}\end{aligned}$$

3.2 Marginal Density Functions

From the expression in (C.11), a density function can be deduced for the coordinate function x_2 (see Theorem 3.1 below). However, first a set of fixed switching probabilities corresponding to the constant vector fields are given.

Assume that the probability of switching from one state to another is given by

$$\begin{aligned}P_{11}(x_2) = 1, \quad P_{21}(x_2) = 1 \quad &\text{for } x_2 < 0 \\ P_{12}(x_2) = 1, \quad P_{22}(x_2) = 1 \quad &\text{for } x_2 > 0 \qquad \text{(C.12)}\end{aligned}$$

where $\{1,2\}$ are discrete states given by $\chi(t)$.

Consider $D_t[\Phi^j](x_2,t)$ from (C.11) in both states with the probabilities given by (C.12),

$$\begin{aligned}D_t[\Phi^-](x_2,t) = &- v_2^- D_{x_2}[\Phi^-](x_2,t) - \lambda\Phi^-(x_2,t) \\ &+ \lambda P_{11}(x_2)\Phi^-(x_2,t) + \lambda P_{21}(x_2)\Phi^+(x_2,t) \\ D_t[\Phi^+](x_2,t) = &- v_2^+ D_{x_2}[\Phi^+](x_2,t) - \lambda\Phi^+(x_2,t) \\ &+ \lambda P_{12}(x_2)\Phi^-(x_2,t) + \lambda P_{22}(x_2)\Phi^+(x_2,t) \,.\end{aligned}$$

In the following, we search for a stationary solution such that $D_t[\Phi^j](x_2,t) = 0$ and

$$v_2^- D_{x_2}[\Phi^-](x_2) = -\lambda\Phi^-(x_2) + \lambda P_{11}(x_2)\Phi^-(x_2) + \lambda P_{21}(x_2)\Phi^+(x_2)\,,$$
$$v_2^+ D_{x_2}[\Phi^+](x_2) = -\lambda\Phi^+(x_2) + \lambda P_{12}(x_2)\Phi^-(x_2) + \lambda P_{22}(x_2)\Phi^+(x_2)\,.$$

For $x_2 < 0$ this implies that

$$v_2^- D_{x_2}[\Phi^-](x_2) = \lambda\Phi^+(x_2),\;\; v_2^+ D_{x_2}[\Phi^+](x_2) = -\lambda\Phi^+(x_2)\,,$$

and for $x_2 > 0$

$$v_2^- D_{x_2}[\Phi^-](x_2) = -\lambda\Phi^-(x_2),\;\; v_2^+ D_{x_2}[\Phi^+](x_2) = \lambda\Phi^-(x_2)\,.$$

The above expressions are ordinary differential equations, which for $x_2 < 0$ have the solutions

$$\Phi^+(x_2) = K_2 \exp\left(\frac{-\lambda x_2}{v_2^+}\right) \tag{C.13}$$

$$\Phi^-(x_2) = -K_2\frac{v_2^+}{v_2^-}\exp\left(\frac{-\lambda x_2}{v_2^+}\right)\,, \tag{C.14}$$

and for $x_2 > 0$

$$\Phi^-(x_2) = K_1 \exp\left(\frac{-\lambda x_2}{v_2^-}\right) \tag{C.15}$$

$$\Phi^+(x_2) = -K_1\frac{v_2^-}{v_2^+}\exp\left(\frac{-\lambda x_2}{v_2^-}\right) \tag{C.16}$$

where K_1 and K_2 are appropriate positive constants.

Continuity

Observe that

$$\phi_{x_2}(x_2) = \Phi^-(x_2) + \Phi^+(x_2)\,. \tag{C.17}$$

It is preferable that $\lim_{x_2\nearrow 0}\phi_{x_2}(x_2) = \lim_{x_2\searrow 0}\phi_{x_2}(x_2)$ which requires that this property holds for $\Phi^-(x_2)$ and $\Phi^+(x_2)$. The limits of these functions are

$$\lim_{x_2\nearrow 0}\Phi^- = -K_2\frac{v_2^+}{v_2^-}, \qquad \lim_{x_2\searrow 0}\Phi^- = K_1$$

and

$$\lim_{x_2 \nearrow 0} \Phi^+ = K_2, \qquad \lim_{x_2 \searrow 0} \Phi^+ = -K_1 \frac{v_2^-}{v_2^+} .$$

Hence, it is required that

$$K_1 \quad = \quad -K_2 \frac{v_2^+}{v_2^-} \tag{C.18}$$

or similar $K_2 = -K_1 \frac{v_2^-}{v_2^+}$.

The integral of the marginal density function along the entire x_2-axis is one. From this, the value of the constant K_1 can be determined;

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \Phi^-(x_2) + \Phi^+(x_2) dx_2 \\ &= \int_{-\infty}^{0} K_1 \exp\left(\frac{-\lambda x_2}{v_2^+}\right) - K_1 \frac{v_2^-}{v_2^+} \exp\left(\frac{-\lambda x_2}{v_2^+}\right) dx_2 \\ &\quad + \int_{0}^{\infty} K_1 \exp\left(\frac{-\lambda x_2}{v_2^-}\right) - K_1 \frac{v_2^-}{v_2^+} \exp\left(\frac{-\lambda x_2}{v_2^-}\right) dx_2 \\ &= K_1 \left(1 - \frac{v_2^-}{v_2^+}\right) \frac{v_2^- - v_2^+}{\lambda} . \end{aligned}$$

Therefore,

$$K_1 = \frac{v_2^+}{v_2^+ - v_2^-} \frac{\lambda}{v_2^- - v_2^+} = \frac{-\lambda v_2^+}{\left(v_2^+ - v_2^-\right)^2} > 0 . \tag{C.19}$$

The function $\Phi^j(x_2, t) = Q^j(x_2, t)\phi_{x_2}(x_2, t)$ for $j \in \{-, +\}$ is a probabilistic function, which can be interpreted as a scaled positional density function. In the case where $\phi_{x_2}(x_2, t) = \phi_{x_2}(x_2)$ is stationary, the parameters $Q^j(x_2, t) = Q^j$ for $j \in \{-, +\}$ are stationary state probabilities. Therefore,

$$\phi_{x_2}(x_2) = (Q^- + Q^+)\phi_{x_2}(x_2) .$$

To determine the constants Q^-, Q^+ consider

$$\Phi^-(x_2) = Q^- \phi_{x_2}(x_2) . \tag{C.20}$$

By substituting (C.13), (C.14) and (C.18) into (C.20) for $x_2 < 0$

$$K_1 \exp\left(\frac{-\lambda x_2}{v_2^+}\right) = Q^- K_1 \left(1 - \frac{v_2^-}{v_2^+}\right) \exp\left(\frac{-\lambda x_2}{v_2^+}\right) ,$$

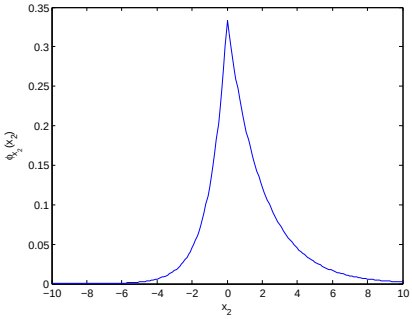

Figure C.5: The stationary marginal density function $\phi_{x_2}(x_2)$ where $\mathbf{v}^- = (1,2)$ and $\mathbf{v}^+ = (1,-1)$.

and $Q^- = \frac{v_2^+}{v_2^+ - v_2^-}$. Since $Q^- + Q^+ = 1$ it follows that $Q^+ = \frac{-v_2^-}{v_2^+ - v_2^-}$. For $x_2 > 0$ the same values of Q^- and Q^+ are obtained.

The above deductions result in the following theorem.

Theorem 3.1. *For Model 1 and the corresponding positional probability given in* (C.7)*, assume that a stationary density function $\phi_{x_2}(x_2)$ exists. Furthermore, assume that the probabilities of switching between the two vector fields are given by* (C.12)*. Then, the stationary density function for the x_2-coordinate function is given by*

$$\phi_{x_2}(x_2) = \frac{\lambda}{v_2^- - v_2^+}\left(\mathbb{I}_{(x_2<0)}e^{-\frac{\lambda}{v_2^+}x_2} + \mathbb{I}_{(x_2>0)}e^{-\frac{\lambda}{v_2^-}x_2}\right) \tag{C.21}$$

where the indicator function is defined

$$\mathbb{I}_{(y>a)} = \begin{cases} 1 & \text{if } y > a \\ 0 & \text{else} \end{cases} .$$

Proof. The expression follows from the above deductions by substituting (C.13), (C.14), (C.15), (C.16) and (C.19) into (C.17). □

We remark that a crucial assumption in the above theorem is the existence of a stationary density function. However, the cyclic behaviour of the system and the fact that the vector fields are constant support the reliability of this assumption.

In Figure C.5 the stationary marginal density function for x_2 is plotted where the parameters are $\mathbf{v}^- = (1,2)$, $\mathbf{v}^+ = (1,-1)$ and $\lambda = 1$. The graph of the stationary density function illustrates that the effect of $|v_2^-| = 2 > 1 = |v_2^+|$ implies a skewness of the probability mass. This follows also theoretically due to the stationary mean and variance, which are given by

$$E[x_2] = \frac{v_2^- + v_2^+}{\lambda} \quad \text{and} \quad \text{Var}[x_2] = \frac{(v_2^-)^2 + (v_2^+)^2}{\lambda^2} .$$

The mean of the stationary density function given in Figure C.5 is $E[x_2] = \frac{2-1}{1} = 1$ and, therefore, the graph illustrates that a bigger amount of probability mass is concentrated on the positive real line than on the negative real line.

Another observation, which is evident from the stationary distribution, is the convergence towards the Filippov solution as the intensity of the Poisson point process increases, that is, as $\lambda \to \infty$. From a model perspective, increasing the intensity corresponds to increasing the number of observations of the current position of the trajectory. Furthermore, by definition the Filippov solution corresponds exactly to the behaviour followed by instantly switching between subsystems around a switching surface.

In subsection 2.2, the relation between η_1 and x_2 was deduced to be $x_2 = \frac{1}{\kappa}\eta_1$. Since the marginal density function for x_2 is known, this relationship gives the marginal density function for η_1.

Corollary 3.2. *Under the assumptions of Theorem 3.1 the marginal density function for $\eta_1 = x_1 - C_1 t$ is given by*

$$\phi_{\eta_1}(\eta_1) = \frac{1}{\kappa}\phi_{x_2}\left(\frac{1}{\kappa}\eta_1\right)$$

where $\phi_{x_2}(\cdot)$ is given in (C.21).

Another consequence of the linear relation $x_2 = \eta_2 = \frac{1}{\kappa}\eta_1$ is that the probability of the event $(\eta_1, \eta_2) \in B$ for $B \subset \mathbb{R}^2$ is concentrated on a straight line. The corresponding joint probability distribution can be expressed explicitly by the density function for $x_2 = \eta_2$.

Corollary 3.3. *Let $L = \{(\eta_1, \eta_2) \mid \eta_1 = \kappa\eta_2\}$ and assume that a stationary density function for η_2 exists. The joint probability distribution of (η_1, η_2) is given by the joint probabilities*

$$P((\eta_1, \eta_2) \in B) = \int_{(B\cap L)_2} \phi_{x_2}(\eta_2)d\eta_2 \quad \text{for } B \subset \mathbb{R}^2 \tag{C.22}$$

where $\phi_{x_2}(\cdot)$ is given in (C.21) *and $(B \cap L)_2$ is the projection of $B \cap L$ on the η_2-axis.*

Proof. The probability that $(\eta_1, \eta_2) \in B$ is

$$P((\eta_1, \eta_2) \in B) \quad = \quad P((\eta_1, \eta_2) \in B \cap L)\ .$$

Observe that the projection of B on the η_2-axis is

$$\begin{aligned}(B \cap L)_2 &= \{\eta_2 | (\eta_1, \eta_2) \in B \cap L\} \\ &= \{\eta_2 | \eta_1 = \kappa\eta_2 \wedge (\eta_1, \eta_2) \in B\}\ .\end{aligned}$$

Therefore,

$$\begin{aligned} P((\eta_1,\eta_2)\in B\cap L) &= P(\eta_1=\kappa\eta_2\wedge(\eta_1,\eta_2)\in B) \\ &= P(\eta_2\in(B\cap L)_2) \\ &= \int_{(B\cap L)_2}\phi_{x_2}(\eta_2)d\eta_2\,. \end{aligned}$$

We remark that the expression in (C.22) implies that no joint density function exists for (η_1,η_2).

4 Perspectives

In this paper, switching dynamics and sliding mode behaviour are investigated under the assumption that the randomized switching is driven by the Poisson point process. This is done in a two-dimensional framework. As mentioned in the introduction, this can easily be generalized to the n-dimensional setting by choosing $x_1\in\mathbb{R}^{n-1}$ and $x_2\in\mathbb{R}$. However, this extension does not provide additional information and, hence, it is left to the reader.

Future perspectives of the presented work are to investigate patterns in the switches. One possible approach is to consider auto- and cross correlation functions as well as their corresponding Fourier transforms, i.e. *power spectra*. The relevance of power spectra from switched control systems is warranted from a noise perspective, i.e. the physical manifestation of control residuals may come as electrical signals for which noise mediation highly depends on the power spectrum of the noise.

An additional step to take is to consider non-constant vector fields allowing a significantly larger validity domain of the provided results.

References

[1] H. K. Khalil, *Nonlinear Systems*, Third, Ed. Prentice-Hall. Inc., 2000.

[2] D. Liberzon, *Switching in systems and control*, ser. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2003. [Online]. Available: http://dx.doi.org/10.1007/978-1-4612-0017-8

[3] V. I. Utkin, J. Guldner, and J. Shi, *Sliding mode control in electromechanical systems*, ser. The Taylor & Francis systems and control book series. London, Philadelphia, PA: Taylor & Francis, 1999.

[4] M. Bhavea, S. Janardhananb, and L. Dewana, "A finite-time convergent sliding mode control for rigid underactuated robotic manipulator," *Systems Science & Control Engineering: An Open Access Journal*, vol. 2, no. 8, pp. 493–499, 2014.

[5] S. Chanda and P. Gogoi, "Trajectory tracking of a 2-link robotic manipulator using adaptive terminal sliding mode controller," *International Journal of Scientific Engineering and Research (IJSER)*, vol. 2, 2014.

[6] A. Sharkawy and S. Salman, "An adaptive fuzzy sliding mode control scheme for robotic systems," *Intelligent Control and Automation*, vol. 2, no. 4, pp. 299–309, 2014.

[7] A. F. Filippov, *Differential equations with discontinuous righthand sides*, ser. Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht, 1988, vol. 18, translated from the Russian. [Online]. Available: http://dx.doi.org/10.1007/978-94-015-7793-9

[8] E. A. Asarin and R. N. Izmaĭlov, "Determination of the sliding speed on a discontinuity surface," *Avtomat. i Telemekh. (transl.,Automation and Remote Control 50(1989),pp.1181-1185)*, no. 9, pp. 43–48, 1989.

[9] M. L. Bujorianu and J. Lygeros, "Toward a general theory of stochastic hybrid systems," in *Stochastic hybrid systems*, ser. Lecture Notes in Control and Inform. Sci. Berlin: Springer, 2006, vol. 337, pp. 3–30.

[10] J. Liu and A. R. Teel, "Generalized solutions to hybrid systems with delays," *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pp. 6169–6174, 2012. [Online]. Available: http://dx.doi.org/10.1109/CDC.2012.6425940

[11] J. Leth, J. G. Rasmussen, H. Schioler, and R. Wisniewski, "A class of stochastic hybrid systems with state-dependent switching noise," *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pp. 4737–4744, 2012. [Online]. Available: http://dx.doi.org/10.1109/CDC.2012.6427010

[12] K. Ito, "On stochastic differential equations," *Mem. Amer. Math. Soc.*, vol. 1951, no. 4, p. 51, 1951.

[13] A. J. Veretennikov, "On strong solutions and explicit formulas for solutions of stochastic integral equations," *Math. USSR Sb.*, vol. 39, no. 3, pp. 387–403, 1981. [Online]. Available: http://dx.doi.org/10.1070/SM1981v039n03ABEH001522

[14] I. Gyöngy and N. Krylov, "Existence of strong solutions for Itô's stochastic equations via approximations," *Probab. Theory Related Fields*, vol. 105, no. 2, pp. 143–158, 1996. [Online]. Available: http://dx.doi.org/10.1007/BF01203833

Paper D

Investigations of the Switching Dynamics in Sliding Mode Control

Maria Simonsen, Henrik Schiøler and John Leth

The paper is accepted for publication in the
proceedings of the 20th World Congress of the International Federation of Automatic Control, 2017.

Abstract

In this paper, a (switching) sliding mode controller, applied to a mechanical system with additive white process noise, is investigated. The practical relevance of this study is a statistical characterization of system performance in terms of the stationary variance of the control error.

The system is modelled with a two-dimensional stochastic differential equation, whose coordinate functions to an extend are analysed separately. In order to determine the stationary variance of one of the coordinate functions, the auto-correlation function for the other coordinate function is approximated with a Fourier series. Finally, analytical results are compared to results from Euler-Maruyama simulation over a wide range of model parameter settings.

Keywords: Stochastic differential equation, Sliding mode, Euler-Maruyama simulation, Switching dynamics, Fokker Planck equation.

1 Introduction

Sliding mode control is a nonlinear control method which typically applies a discontinuous control signal to force a system to behave according to prescribed closed loop dynamics. Essentially, the control procedure consists of two parts: Firstly, the controller forces the system state to approach a so called sliding surface and, secondly, to slide along the surface towards the operating point. Near the operating point, the main purpose of the sliding mode controller is to keep this position and respond accordingly to any external disturbance (noise) which affects the system. The sliding surface is found as the sub-manifold of the desired closed loop dynamics. Robustness to external disturbances is achieved by the design of a feedback control, which is discontinuous across the sliding surface. The discontinuity creates in practice rapid switching and in theory additional challenges w.r.t. e.g. existence and uniqueness of solutions to model equations.

The discontinuity induced by the controller brings the main challenges in the analysis of the system and is a main motivation behind the investigations of the switching dynamics. A solid amount of literature exists on the application and analysis of sliding mode control. Among others, see [1, 2]. Recent papers on application of sliding mode control are [3, 4].

In this paper, the system is modelled with stochastic differential equations (SDEs). Solutions to the SDE's are then considered by using appropriate approximations of practical implementations of switching, where the latter may include various imprecisions such as delay, hysteresis and continuous approximation of switching discontinuity. As a result, the system is represented with a two-dimensional SDE with discontinuous drift coefficient and constant diffusion coefficient.

Solutions to SDEs have over a long period been a subject of great interest, both in the form of existence, uniqueness, explicit closed form solutions and construction of numerical approximations, see [5–7]. Whereas existence and uniqueness are established for the discontinuous bounded drift case (see [8, 9]), it is not yet proven for SDEs with unbounded discontinuous drift coefficients. Neither specific characteristics such as transient and stationary distributions, nor auto- and cross-covariance characteristics have been established for SDEs with discontinuous drift coefficient.

In order to apply some of the known results on SDEs with discontinuous bounded drift, a coordinate transformation of the system is introduced. This transformation implies that the discontinuous dynamics is isolated to only one of the coordinate functions of the two-dimensional SDE, which, additionally, has bounded drift coefficient. By this approach we are able to initiate the study of the two-dimensional SDE through well-known existence results of one of its coordinate functions. We apply the theory of Fourier series to the Fokker-Planck equation for the SDE with discontinuous bounded drift coefficient [10, 11]. This gives, by approximation methods, a time-dependent conditional density function from which an estimate on the auto-correlation is deduced. Finally, by including the auto-correlation function into the analysis of the variance, an estimate of the stationary variance is determined.

Notation: Let $x : \Omega \times \mathbb{R}^+ \to \mathbb{S}$ denote a (stochastic) process on a probability space (Ω, Σ, P) with values in the state space $\mathbb{S}$. Throughout this paper we suppress the processes dependency on the variable from the measure space (Ω, Σ), that is, we write $x(t)$ in place of $x(\cdot, t)$.

The outline of the paper is as follows: In section 2 we provide the definition of the system, which is subject to subsequent analysis. In section 3 we initiate the analysis of system (D.3) through a coordinate transformation producing new coordinate functions η_1 and η_2. In section 4 the η_1 coordinate function is analysed in order to give a bound on its variance. To estimate the variance of η_1, an estimate of the autocorrelation function of η_2 is needed, which is presented in section 5 and 6. Finally, the main result is presented in section 7. Numerical results are presented in sections 8 and compared with results generated by Euler-Maruyama simulations.

2 System Definition

Consider the idealized model of a mechanical system (with mass 1) given by the (control) system of first order differential equations

$$\dot{x}_1 = x_2 \tag{D.1a}$$

$$\dot{x}_2 = F(x, u) \tag{D.1b}$$

where $x_1 : [0, T] \to \mathbb{R}$ is the position at time $t \in [0, T]$, $x = (x_1, x_2) : [0, T] \to \mathbb{R}^2$, u is the control variable and $F : \mathbb{R}^2 \to \mathbb{R}$ comprises control forces, conservative

forces, friction forces and other deterministic external forces.

However, any realistic model describing the behaviour of a mechanical system should include the influence of non-deterministic (random) forces. Such forces are often modelled by means of the Wiener process (appearing as white noise in some expositions). Therefore, let $W : [0, T] \to \mathbb{R}$ (with $W(t) = W_t$) denote a real valued standard Wiener process with initial value $W_0 = 0$ and consider, instead of (D.1), the (control) system of SDEs

$$dx_1 = x_2 dt \tag{D.2a}$$
$$dx_2 = F(x, u)dt + dW_t \, . \tag{D.2b}$$

In the sequel, we study the stochastic behaviour of the SDE given in (D.2) when a sliding mode controller is applied. The analysis is limited to the case where the force F comprises only friction force modelled as $-\alpha x_2$ where $\alpha > 0$ is a viscous friction coefficient as well as feedback control forces. In summary, we consider the system

$$dx_1 = x_2 dt \tag{D.3a}$$
$$dx_2 = -\alpha x_2 dt + u dt + dW_t \, . \tag{D.3b}$$

The control u is designed as a sliding mode controller with switching across a sliding surface

$$S = \{(x_1, x_2) : a x_1 + x_2 = 0\}$$

where the design constant a is chosen to ensure S to be a stable manifold. To ensure reaching the surface and maintain sliding, the control u is designed as

$$u = -k \operatorname{sgn}(a x_1 + x_2),$$

with k being a constant gain. Next, we fix $a = \alpha$. This brings useful properties to the dynamics of the system which is advantageous in the following analysis of the system. However, there is a tradeoff as this implies that the ability to influence the control variable u is restricted to the constant gain k.

3 Preliminary Analysis of the System

In system D.3, the control function u contributes with discontinuous dynamics driven by both coordinate functions. In order to simplify this challenge, we introduce a coordinate transformation of the system.

Let a (linear) coordinate transformation from (x_1, x_2) coordinates to (η_1, η_2) coordinates be defined by

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{bmatrix} 1 & -\alpha \\ \alpha & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} . \tag{D.4}$$

The coordinates are defined such that the η_1-axis is parallel with the sliding surface and the η_2-axis is perpendicular to the sliding surface.

The new coordinate system D.3 then reads

$$\begin{aligned} d\eta_1 &= dx_1 - \alpha dx_2 \\ &= (-\alpha\eta_1 + \eta_2)dt - \alpha u dt - \alpha dW_t \qquad \text{(D.5a)} \\ d\eta_2 &= \alpha dx_1 + dx_2 = u dt + dW_t \qquad \text{(D.5b)} \end{aligned}$$

where $u : [0, T] \to \mathbb{R}$ is the control signal given by

$$u = -k\,\mathrm{sgn}(\eta_2),$$

with k being a constant gain.

3.1 The η_2 Coordinate Function

The SDE given in (D.5b) depends only on η_2 and the Wiener process W_t and, therefore, it can be treated as an independent one-dimensional SDE. The drift coefficient is bounded and discontinuous and the diffusion coefficient is equal to the identity. For such an SDE, it is proven by A. J. Veretennikov that a strongly unique solution exists [9]. Furthermore, for the particular discontinuous SDE in (D.5b) some additional information has previously been obtained.

Firstly, the stationary density function is known (see paper A), and is given by

$$\Phi_{\eta_2}(\eta_2) = \mathbb{I}_{(\eta_2<0)} k e^{2k\eta_2} + \mathbb{I}_{(\eta_2>0)} k e^{-2k\eta_2} \qquad \text{(D.6)}$$

where $\mathbb{I}_{(\cdot)}$ denotes the indicator function. From the stationary density function, the stationary mean and variance are determined to be

$$\mathbb{E}[\eta_2] = 0 \quad \text{and} \quad \mathrm{Var}[\eta_2] = \frac{1}{2k^2}\,. \qquad \text{(D.7)}$$

Secondly, it is proven that numerical solutions produced with the Euler-Maruyama method converge to the strong solution of (D.5b) (see paper B).

However, there are still open questions regarding the auto- and cross-correlation functions for the η_2 coordinate function. The auto-correlation analysis is addressed in Section 6 of this paper.

3.2 The η_1 Coordinate Function

Investigations of the dynamics in the η_1 coordinate function is the main contribution of this paper. More specifically, we will derive an estimate of the system behaviour near the system's operating point based on an analysis of the variance.

The following section initiates this analysis by application of Itô calculus to the η_1 coordinate function.

4 Itô Calculus on the η_1 Coordination Function

In the following, Itô's lemma [6, Thm.4.1.2] is applied to the stochastic process

$$Y(\eta_1, t) = \eta_1 e^{\alpha t}$$

where $Y : \mathbb{R}^2 \to \mathbb{R}; (z,t) \mapsto ze^{\alpha t}$ and η_1 is given as in (D.5a). We remark that the SDE in (D.5a) can be treated independently of (D.5b) in Itô's lemma, as long as the local drift and diffusion coefficient are non-anticipating[1]. For further information see [12, p.102].

Application of Itô's lemma gives

$$dY \quad = \quad e^{\alpha t}\left(\eta_2 - \alpha u\right) dt - \alpha e^{\alpha t} dW_t$$

which, when multiplied with $e^{-\alpha t}$, yields the following integral expression for the η_1 coordinate function

$$\eta_1(t) \quad = \quad \eta_1(0)e^{-\alpha t} + \int_0^t e^{\alpha(s-t)}\left(\eta_2 - \alpha u\right) ds - \int_0^t \alpha e^{\alpha(s-t)} dW_s \,. \quad \text{(D.8)}$$

By a similar procedure, Itô's lemma is applied to $Z(\eta_1, t) = \eta_1^2$ to obtain

$$dZ \quad = \quad \left(-2\alpha\eta_1^2 + 2\eta_1\left(\eta_2 + \alpha k \operatorname{sgn}(\eta_2)\right) + \alpha^2\right) dt - 2\eta_1 \alpha dW_t \,.$$

Next, we substitute the expression of η_1 given in (D.8) into (D.9) to obtain an integral form for $r \in [0, T]$. Then, taking the expectation of $\eta_1(t)$ and $Z(\eta_1, r)$ and thereafter applying Fubini's theorem gives

$$\mathbb{E}[\eta_1] \quad = \quad \eta_1(0)e^{-\alpha t} + \int_0^t e^{\alpha(s-t)}\mathbb{E}\left[\eta_2 - \alpha u\right] ds - \mathbb{E}\left[\int_0^t \alpha e^{\alpha(s-t)} dW_s\right],$$

and

$$\begin{aligned}
\mathbb{E}[Z(\eta_1, r)] = \mathbb{E}[Z(\eta_1, 0)] &+ \int_0^r -2\alpha\mathbb{E}[Z(\eta_1, t)]dt \\
&+ 2\int_0^r \int_0^t e^{\alpha(s-t)}\mathbb{E}\left[\left(\eta_2(t) + \alpha k \operatorname{sgn}(\eta_2(t))\right) \cdot \left(\eta_2(s) + \alpha k \operatorname{sgn}(\eta_2(s))\right)\right] ds dt \\
&- 2\int_0^r \mathbb{E}\left[\int_0^t \alpha e^{\alpha(s-t)}\left(\eta_2(t) + \alpha k \operatorname{sgn}(\eta_2(t))\right) dW_s\right] dt \\
&+ \int_0^r \mathbb{E}[\alpha^2] dt - 2\mathbb{E}\left[\int_0^r \eta_1(t)\alpha dW_t\right] .
\end{aligned}$$

[1] A function $f(s)$ is non-anticipating if it is independent of the behaviour of the Wiener process in the future of s, that is, $f(s)$ is statistically independent of $W_t - W_s$ for all $s < t$

Recall that the stochastic Itô integral $I(f) = \int_0^t f(\cdot,s)dW_s$ is a martingale if $\mathbb{E}[\int_0^t f^2(\cdot,s)ds] < \infty$ and $f(\cdot,s)$ is non-anticipating. Hence, in this case the expectation of $I(f)$ is zero and the last terms in the expression of $\mathbb{E}[\eta_1]$ and $\mathbb{E}[Z(\eta_1,r)]$ vanish. Therefore, the expectation $\mathbb{E}[\eta_1]$ is reduced to

$$\begin{aligned}\mathbb{E}[\eta_1] =& \eta_1(0)e^{-\alpha t} + \int_0^t e^{\alpha(s-t)}\left(\mathbb{E}[\eta_2] + \alpha k \mathbb{E}[\operatorname{sgn}(\eta_2)]\right) ds \\ =& \eta_1(0)e^{-\alpha t},\end{aligned}$$

under stationary assumption on the η_2 coordinate function (see (D.6) and (D.7)).

Furthermore, using the result of the Appendix (see (D.28)) we have

$$\begin{aligned}&\mathbb{E}[Z(\eta_1,r)] \\ &\quad\approx Z(\eta_1,0) + \int_0^r -2\alpha\mathbb{E}[Z(\eta_1,t)]dt + 2\int_0^r\int_0^t e^{\alpha(s-t)}\Big(\mathbb{E}\left[\eta_2(t)\eta_2(s)\right] \\ &\qquad + \alpha k\mathbb{E}[\operatorname{sgn}(\eta_2(t))\eta_2(s)] + \alpha k\mathbb{E}[\eta_2(t)\operatorname{sgn}(\eta_2(s))] \\ &\qquad + \alpha^2k^2\mathbb{E}[\operatorname{sgn}(\eta_2(t))\operatorname{sgn}(\eta_2(s))]\Big)dsdt - 2\int_0^r \frac{\alpha(\alpha k^2+1)}{\alpha+k^2}dt + \int_0^r \alpha^2 dt,\end{aligned}$$

By differentiating observe that

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[Z(\eta_1,t)] \approx& -2\alpha\mathbb{E}[Z(\eta_1,t)] + 2\int_0^t e^{\alpha(s-t)}\Big(\mathbb{E}\left[\eta_2(t)\eta_2(s)\right] \\ &+ \alpha k\mathbb{E}[\operatorname{sgn}(\eta_2(t))\eta_2(s)] + \alpha k\mathbb{E}[\eta_2(t)\operatorname{sgn}(\eta_2(s))] \\ &+ \alpha^2k^2\mathbb{E}[\operatorname{sgn}(\eta_2(t))\operatorname{sgn}(\eta_2(s))]\Big)ds - 2\frac{\alpha(\alpha k^2+1)}{\alpha+k^2} + \alpha^2,\end{aligned} \tag{D.9}$$

and, therefore, an approximated stationary variance of η_1 can be determined from $0 = \frac{d}{dt}\mathbb{E}[Z(\eta_1,t)] = \frac{d}{dt}\mathbb{E}[\eta_1^2]$ whenever the four auto- and cross-covariance terms in the integrand are known.

Since all auto- and cross-covariances only depend on the η_2 coordinate function, the density of η_2 is analysed in the following section. This analysis is based on a Fourier series expansion of the density obtained via the Fokker-Planck equation.

5 A Discrete Fourier Series Representation of the Marginal Density Function for η_2

The Fokker-Planck equation related to $\eta_2(t)$ is

$$\begin{aligned}\frac{\partial}{\partial t}p(\eta_2,t) =&\frac{\partial}{\partial \eta_2}\left[k\,\mathrm{sgn}(\eta_2)p(\eta_2,t)\right]+\frac{1}{2}\frac{\partial^2}{\partial \eta_2^2}p(\eta_2,t)\\ =&k\left[\frac{\partial}{\partial \eta_2}\,\mathrm{sgn}(\eta_2)\right]p(\eta_2,t)+k\,\mathrm{sgn}(\eta_2)\frac{\partial}{\partial \eta_2}p(\eta_2,t)+\frac{1}{2}\frac{\partial^2}{\partial \eta_2^2}p(\eta_2,t)\end{aligned} \tag{D.10}$$

where $p(\eta_2,t)$ is the marginal density function for the η_2 coordinate function. A truncated version of the density function can be represented as a Fourier series over the interval $[-L,L]$,

$$p(\eta_2,t) \;=\; \frac{a(t)}{2}+\sum_{n=1}^{\infty}b_n(t)\cos\left(\frac{\pi n\eta_2}{L}\right)+\sum_{n=1}^{\infty}c_n(t)\sin\left(\frac{\pi n\eta_2}{L}\right) \tag{D.11}$$

where

$$a(t)=\frac{1}{L}\int_{-L}^{L}p(\eta_2,t)d\eta_2\,, \tag{D.12a}$$

$$b_n(t)=\frac{1}{L}\int_{-L}^{L}p(\eta_2,t)\cos\left(\frac{\pi n\eta_2}{L}\right)d\eta_2\,, \tag{D.12b}$$

$$c_n(t)=\frac{1}{L}\int_{-L}^{L}p(\eta_2,t)\sin\left(\frac{\pi n\eta_2}{L}\right)d\eta_2\,. \tag{D.12c}$$

Furthermore, over the interval $[-L,L]$, the sign-function $\mathrm{sgn}(\eta_2)$ and its (distributional) derivative $\frac{\partial}{\partial \eta_2}\mathrm{sgn}(\eta_2)=2\delta$, with δ denoting the Dirac delta distribution, can be represented as the Fourier series

$$\mathrm{sgn}(\eta_2)=\sum_{n=1}^{\infty}\frac{2(1-\cos(\pi n))}{\pi n}\sin\left(\frac{\pi n\eta_2}{L}\right), \tag{D.13}$$

and

$$\frac{\partial}{\partial \eta_2}\mathrm{sgn}(\eta_2)=2\delta(\eta_2)=\frac{1}{L}+\frac{2}{L}\sum_{n=1}^{\infty}\cos\left(\frac{\pi n\eta_2}{L}\right)\,. \tag{D.14}$$

We now outline how the coefficient function

$$b_1(t),c_1(t),b_2(t),c_2(t),\ldots$$

from (D.12) and, therefore, also the density in (D.11), can be approximated as solutions to a system of linear ordinary differential equations. Details are omitted due to limited space.

We write $p(\eta_2, t)$ in terms of its Fourier series representation including the series for $\operatorname{sgn}(\eta_2)$ and $\frac{\partial}{\partial \eta_2}\operatorname{sgn}(\eta_2) = 2\delta(\eta_2)$ in the Fokker-Planck equation (D.10). Then for the series representation we expand the derivatives of the Fokker Planck equation and finally collect a truncated (finite) subset of trigonometric terms to obtain the following linear coefficient dynamics

$$\frac{d}{dt}\begin{pmatrix} z \\ a \end{pmatrix} = \begin{bmatrix} A & b \\ 0 & 0 \end{bmatrix}\begin{pmatrix} z \\ a \end{pmatrix} \tag{D.15}$$

where $A \in \mathbb{R}^{2\ell \times 2\ell}$ (with ℓ representing the number of terms),

$$z = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_\ell(t) \\ c_1(t) \\ c_2(t) \\ \vdots \\ c_\ell(t) \end{pmatrix}, \quad b = \begin{pmatrix} k \\ k \\ \vdots \\ k \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad z(t_0) = \begin{pmatrix} \cos\left(\frac{\pi\eta_2(t_0)}{L}\right) \\ \cos\left(\frac{2\pi\eta_2(t_0)}{L}\right) \\ \vdots \\ \cos\left(\frac{\ell\pi\eta_2(t_0)}{L}\right) \\ \sin\left(\frac{\pi\eta_2(t_0)}{L}\right) \\ \sin\left(\frac{2\pi\eta_2(t_0)}{L}\right) \\ \vdots \\ \sin\left(\frac{\ell\pi\eta_2(t_0)}{L}\right) \end{pmatrix}$$

and $a(t_0) = \frac{1}{L}$. The solution to (D.15) is given explicitly by

$$\begin{pmatrix} z(t) \\ a(t) \end{pmatrix} = \exp\left(\begin{bmatrix} A & b \\ 0 & 0 \end{bmatrix} t\right)\begin{pmatrix} z(t_0) \\ a(t_0) \end{pmatrix},$$

thus determines an approximated conditional density function $\tilde{p}$ given by

$$\begin{aligned}\tilde{p}(\eta_2(t)|\eta_2(t_0)) &= \frac{a}{2} + \sum_{n=1}^{\ell} b_n(t, \eta_2(t_0)) \cos\left(\frac{\pi n \eta_2}{L}\right) \\ &\quad + \sum_{n=1}^{\ell} c_n(t, \eta_2(t_0)) \sin\left(\frac{\pi n \eta_2}{L}\right)\end{aligned}$$

where $\{b_n(t, \eta_2(t_0)), c_n(t, \eta_2(t_0)) : n \in \{1, \ldots, \ell\}\}$ are the approximated coefficients.

In the following section, we return to the main problem to determine the auto- and cross-covariance terms appearing in the expression of $\frac{d}{dt}\mathbb{E}[Z(\eta_1, t)] = \frac{d}{dt}\mathbb{E}[\eta_1^2]$. More precisely, the approximated conditional density function is applied to determine approximations of $\mathbb{E}[\eta_2(t)\eta_2(s)]$, $\mathbb{E}[\eta_2(t)\operatorname{sgn}(\eta_2(s))]$ and $\mathbb{E}[\operatorname{sgn}(\eta_2(t))\operatorname{sgn}(\eta_2(s))]$ for $t, s \in [0, T]$.

6 The Auto-Correlation Function and Related Expressions

The auto-correlation function is given by

$$\begin{aligned}\mathbb{E}[\eta_2(t)\eta_2(s)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\eta_2 y f_{\eta_2(t),\eta_2(s)}(\eta_2,y)d\eta_2 dy \\ &= \int_{-\infty}^{\infty} y f_{\eta_2(s)}(y)\int_{-\infty}^{\infty}\eta_2 f_{\eta_2(t)|\eta_2(2)}(\eta_2|y)d\eta_2 dy \\ &= \int_{-\infty}^{\infty} y f_{\eta_2(s)}(y)\mathbb{E}[\eta_2(t)|y]dy,\end{aligned}$$

where $f_{\eta_2(s)}(y)$ is the density function for the distribution of $\eta_2(s)$, which by assumption is equal to Φ_{η_2} given in (D.6). The density $f_{\eta_2(t)}(\eta_2|y)$ is substituted with $\tilde{p}(\eta_2(t)|y)$, which is the approximated conditional density function given an initial condition $y = \eta_2(s)$. Therefore, the conditional expectation, $\mathbb{E}[\eta_2(t)|y]$ can be approximated by

$$\begin{aligned}\mathbb{E}[\eta_2(t)|y] &\approx \int_{-L}^{L}\eta_2\tilde{p}(\eta_2(t)|y)d\eta_2 \\ &= \int_{-L}^{L}\eta_2\frac{a(t)}{2}d\eta_2 + \sum_{m=1}^{\ell}\int_{-L}^{L}\eta_2 b_m(t,y)\cos\left(\frac{\pi m\eta_2}{L}\right)d\eta_2 \\ &\quad + \sum_{m=1}^{\ell}\int_{-L}^{L}\eta_2 c_m(t,y)\sin\left(\frac{\pi m\eta_2}{L}\right)d\eta_2 \\ &= \sum_{m=1}^{\ell} c_m(t,y)\frac{-2L^2}{\pi m}\cos(\pi m)\ .\end{aligned}$$

From this

$$\mathbb{E}[\eta_2(t)\eta_2(s)] \approx \int_{-\infty}^{\infty} y\Phi_{\eta_2}(y)\sum_{m=1}^{\ell} c_m(t,y)\frac{-2L^2}{\pi m}\cos(\pi m)dy \qquad \text{(D.16)}$$

where Φ_{η_2} is the stationary density function for η_2 given in (D.6). Since coefficients $c_m(t)$ appear as solutions to (D.15), the integrand appears as a weighted sum of products of linear functions and complex exponentials and in that way gives a closed form solution.

A similar procedure can be applied to determine

$$\mathbb{E}[\text{sgn}(\eta_2(t))\eta_2(s)]$$

for $s \leq t$. Observe that

$$\mathbb{E}[\mathrm{sgn}(\eta_2(t))\eta_2(s)] \quad = \quad \int_{-\infty}^{\infty} y\mathbb{E}[\mathrm{sgn}(\eta_2(t))|y]\Phi_{\eta_2}(y)dy \; .$$

The expectation $\mathbb{E}[\mathrm{sgn}(\eta_2(t))|y]$ is given by

$$\mathbb{E}[\mathrm{sgn}(\eta_2(t))|y] \quad = \quad \int_{-\infty}^{\infty} \mathrm{sgn}(\eta_2) f_{\eta_2(t)|y}(\eta_2)d\eta_2 \; ,$$

and can be approximated by

$$\begin{aligned}
\mathbb{E}[\mathrm{sgn}(\eta_2(t))|y] &\approx \int_{-L}^{L} \mathrm{sgn}(\eta_2)\tilde{p}(\eta_2(t)|y)d\eta_2 \\
&= \int_{-L}^{L} \mathrm{sgn}(\eta_2)\frac{a}{2}d\eta_2 + \sum_{m=1}^{\ell}\int_{-L}^{L} \mathrm{sgn}(\eta_2)b_m(t,y)\cos\left(\frac{\pi m\eta_2}{L}\right)d\eta_2 \\
&\quad + \sum_{m=1}^{\ell}\int_{-L}^{L} \mathrm{sgn}(\eta_2)c_m(t,y)\sin\left(\frac{\pi m\eta_2}{L}\right)d\eta_2 \\
&= \sum_{m=1}^{\ell} c_m(t,y)\frac{2L}{\pi m}(1-\cos(\pi m)) \; .
\end{aligned}$$

Hence

$$\mathbb{E}[\mathrm{sgn}(\eta_2(t))\eta_2(s)] \approx \int_{-\infty}^{\infty} y\Phi_{\eta_2}(y)\sum_{m=1}^{\ell} c_m(t,y)\frac{2L}{\pi m}(1-\cos(\pi m))dy \; . \quad \text{(D.17)}$$

By the same procedure we obtain

$$\mathbb{E}[\eta_2(t)\,\mathrm{sgn}(\eta_2(s))] \;\approx\; \int_{-\infty}^{\infty} \mathrm{sgn}(y)\Phi_{\eta_2}(y)\sum_{m=1}^{\ell} c_m(t,y)\frac{-2L^2}{\pi m}\cos(\pi m)dy \;\; , \quad \text{(D.18)}$$

and

$$\begin{aligned}
&\mathbb{E}[\mathrm{sgn}(\eta_2(t))\,\mathrm{sgn}(\eta_2(s))] \\
&\qquad \approx \int_{-\infty}^{\infty} \mathrm{sgn}(y)\Phi_{\eta_2}(y)\sum_{m=1}^{\ell} c_m(t,y)\frac{2L}{\pi m}(1-\cos(\pi m))dy \; . \quad \text{(D.19)}
\end{aligned}$$

In the sequel, the approximations of the expectations are applied to determine an upper bound on the stationary variance of η_1.

7 The Main Result

For $s < t$ let $C(t-s)$ be the sum of the auto- and cross-covariance terms appearing in (D.9), that is

$$C(t-s) = \mathbb{E}\left[\eta_2(t)\eta_2(s)\right] + \alpha k \mathbb{E}[\operatorname{sgn}(\eta_2(t))\eta_2(s)] + \alpha k \mathbb{E}[\eta_2(t)\operatorname{sgn}(\eta_2(s))] + \alpha^2 k^2 \mathbb{E}[\operatorname{sgn}(\eta_2(t))\operatorname{sgn}(\eta_2(s))].$$

Notice that $C(t-s)$ can be approximated by the sum of (D.16)-(D.19) and by this approximation $C(t-s)$ is bounded.

Substituting $C(t-s)$ into (D.9) yields

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[\eta_1^2] &\approx -2\alpha\mathbb{E}[\eta_1^2] - 2\frac{\alpha(\alpha k^2+1)}{\alpha+k^2} + \alpha^2 + 2\int_0^t e^{\alpha(s-t)}C(t-s)ds \\ &\approx -2\alpha\mathbb{E}[\eta_1^2] - 2\frac{\alpha(\alpha k^2+1)}{\alpha+k^2} + \alpha^2 + g(t)\end{aligned}$$

where $g(t)$ denotes the integral term. Since boundedness of C implies a well defined limit, g_∞, of $g(t)$ for $t \to \infty$, the approximated stationary variance of η_1 can be determined as

$$\mathbb{E}[\eta_1^2] \approx -\frac{\alpha k^2+1}{\alpha+k^2} + \frac{\alpha}{2} + \frac{g_\infty}{2\alpha}\,. \tag{D.20}$$

7.1 The Original Coordinates

The analysis above also gives approximated stationary variances of the original coordinates. From (D.4) we get

$$\begin{aligned}x_1 &= \frac{1}{1+\alpha^2}\eta_1 + \frac{\alpha}{1+\alpha^2}\eta_2 \\ x_2 &= -\frac{\alpha}{1+\alpha^2}\eta_1 + \frac{1}{1+\alpha^2}\eta_2\,,\end{aligned}$$

implying

$$\operatorname{Var}[x_1] = \frac{1}{(1+\alpha^2)^2}\operatorname{Var}[\eta_1] + \frac{\alpha^2}{(1+\alpha^2)^2}\operatorname{Var}[\eta_2] + \frac{2\alpha}{(1+\alpha^2)^2}\mathbb{E}[\eta_1\eta_2]\,,$$

with a similar expression for $\operatorname{Var}[x_2]$. The first two terms are given in (D.20) and (D.7). The last term can be approximated by identifying the term $\eta_1\eta_2$ from (D.9), and then using (D.16) and (D.18). More precisely

$$\mathbb{E}[\eta_1\eta_2] \approx \lim_{t\to\infty}\int_0^t e^{\alpha(s-t)}\tilde{C}(t-s)ds$$

$k\backslash\alpha$	0.1	0.3	0.5	0.7
0.1	(4438, 49.82)	(538.43, 50.27)	(209.09, 53.31)	(99.76, 49.82)
0.3	(322.30, 5.68)	(50.14, 5.84)	(20.00, 5.72)	(10.77, 5.68)
0.5	(68.95, 2.08)	(13.02, 2.08)	(5.72, 2.03)	(3.47, 2.08)
0.7	(20.19, 1.01)	(4.90, 1.01)	(2.33, 1.04)	(1.53, 1.08)

Table D.1: The (Euler-Maruyama) estimated variance of $(\eta_1(50), \eta_2(50))$ determined for different values of k and α.

$k\backslash\alpha$	0.1	0.3	0.5	0.7
0.1	(4482, 50)	(531.5, 50)	(194.3, 50)	(100.0, 50)
0.3	(324.96, 5.56)	(49.95, 5.56)	(19.73, 5.56)	(10.68, 5.56)
0.5	(70.02, 2)	(13.20, 2)	(7.5, 2)	(3.44, 2)
0.7	(23.30, 1.02)	(5.18, 1.02)	(2.47, 1.02)	(1.57, 1.02)

Table D.2: The stationary variance of (η_1, η_2) calculated for different values of k and α.

with $\tilde{C}(t-s)$ approximated by the sum of (D.16) and (D.18), that is

$$\tilde{C}(t-s) = \mathbb{E}\left[\eta_2(t)\eta_2(s)\right] + \alpha k \mathbb{E}[\eta_2(t)\,\mathrm{sgn}(\eta_2(s))]\,.$$

In the following section, system (D.5) is simulated with the Euler-Maruyama method to illustrate the validity of the method.

8 Simulation

This section presents a comparison between the stationary variance (D.20) and an estimated variance determined from simulations via the Euler-Maruyama method.

The (η_1, η_2) coordinate functions are simulated with the Euler-Maruyama method with step size $h = 2^{-10}$ for $t \in (0, 50)$. For fixed value of k and α, the simulations are repeated 4000 times and the variance of the end-points of the realizations is determined. The estimated variances are presented in Table D.1. The stationary variances given by (D.20) and (D.7) are presented in Table D.2 for comparison.

Similar, from section 7.1 and the values in Table D.1 and D.2 we also get Table D.3 and D.4 comprising estimated and stationary variances of the original x_1 coordinate.

Finally, Figures D.1 and D.2 show simulated trajectories for η_1, η_2 and x_1, x_2 respectively and a variation of values of k and α. Initial values are in all cases

$k\backslash\alpha$	0.1	0.3	0.5	0.7
0.1	4439	538.74	209.06	100.00
0.3	322.20	50.17	19.78	10.45
0.5	68.97	13.00	5.59	3.22
0.7	20.20	4.84	2.21	1.36

Table D.3: The (Euler-Maruyama) estimated variance of $x_1(50)$ determined for different values of k and α.

$k\backslash\alpha$	0.1	0.3	0.5	0.7
0.1	4483	531.9	194.7	100.3
0.3	325.03	50.21	20.04	10.89
0.5	70.02	13.69	6.13	3.63
0.7	23.32	5.30	2.65	1.72

Table D.4: The stationary variance of x_1 calculated for different values of k and α.

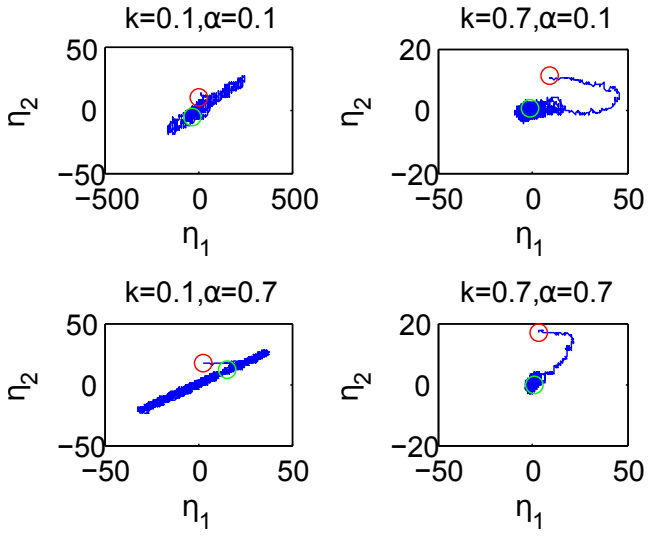

Figure D.1: Simulated η_1, η_2. Initial values shown by red circle and final values shown by green circle.

$(x_1(0), x_2(0)) = (10, 10)$. Notice that for $k = 0.1$ and $\alpha = 0.1$ the reaching and sliding behaviour are hidden in large fluctuations, whereas for the remaining cases reaching and sliding phases are observable.

9 Perspectives

In this paper, we give an estimate of the control error which is induced by the application of a sliding mode controller to a mechanical system. The significance of this result is rooted in the attempt to describe (or estimate) the collected behaviour of a family of physical systems connected via certain switching laws, which are governed by some kind of sliding mode control.

The presented results are for a two-dimensional system with the control input isolated to one coordinate function. Thereby, it is possible to introduce a coordinate transformation which reduces the complexity of the discontinuous dynamics provided

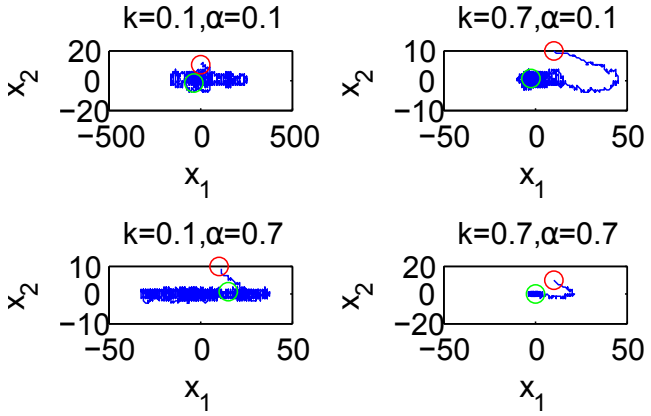

Figure D.2: Simulated x_1, x_2. Initial values shown by red circle and final values shown by green circle.

by the sliding mode controller. For systems of higher dimension with the same structure, i.e. system that can be formulated in the Phase Variable Form, a corresponding coordinate transformation will result in a similar simplification of the discontinuous challenges. Thus, by the same procedure presented in this paper, an estimate of the system behaviour can be evaluated.

Appendix

Consider the integral term appearing in the expression of $\mathbb{E}[Z(\eta_1, r)]$

$$\begin{aligned}\int_0^t \alpha e^{\alpha(s-t)} \left(\eta_2(t) + \alpha k \operatorname{sgn}(\eta_2(t))\right) dW_s \\ = \alpha k \operatorname{sgn}(\eta_2(t)) x(t) + \eta_2(t) x(t)\end{aligned} \tag{D.21}$$

where the process x is defined by

$$x(t) = \int_0^t \alpha e^{\alpha(s-t)} dW_s \, .$$

It is easily recognized that x is the solution of the following SDE

$$dx = -\alpha x dt + \alpha dW_t, \quad x(0) = 0 \, . \tag{D.22}$$

Let the functions $F, Q : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$F(z) = z_1 z_2 \quad \text{and} \quad Q(z) = z_1 \operatorname{sgn}(z_2) \, .$$

The gradients and Hessians are

$$F_z(z) = \begin{bmatrix} z_2 \\ z_1 \end{bmatrix} \text{ and } Q_z(z) = \begin{bmatrix} \operatorname{sgn}(z_2) \\ z_1\delta(z_2) \end{bmatrix},$$

$$F_{zz}(z) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } Q_{zz}(z) = \begin{bmatrix} 0 & \delta(z_2) \\ \delta(z_2) & z_1\delta'(z_2) \end{bmatrix},$$

with δ' denoting the distributional derivative of the Dirac delta distribution.

Using (D.5b) and (D.22) we see that the process z defined by $z = \begin{bmatrix} x \\ \eta_2 \end{bmatrix}$, satisfies the SDE

$$dz = \mu dt + GdW_t = \begin{bmatrix} -\alpha x \\ -k\operatorname{sgn}(\eta_2) \end{bmatrix} dt + \begin{bmatrix} \alpha \\ 1 \end{bmatrix} dW_t .$$

Hence, by Itô's lemma

$$\begin{aligned} dF &= (F_z(x,\eta_2)^T\mu + \frac{1}{2}tr(G^T F_{zz}(x,\eta_2)G))dt + F_z(x,\eta_2)^T GdW_t \\ &= (-\alpha F - kQ + \alpha)dt + F_z(x,\eta_2)^T GdW_t , \end{aligned}$$

and

$$\begin{aligned} dQ &= (Q_z(x,\eta_2)^T\mu + \frac{1}{2}tr(G^T Q_{zz}(x,\eta_2)G))dt + Q_z(x,\eta_2)^T GdW_t \\ &= (-(\alpha + k\delta(\eta_2))Q + \frac{1}{2}(2\alpha\delta(\eta_2) + x\delta'(\eta_2)))dt + Q_z(x,\eta_2)^T GdW_t . \end{aligned}$$

Taking expectations, differentiating and solving for stationarity yields

$$\mathbb{E}[F] = \frac{\alpha - k\mathbb{E}[Q]}{\alpha} , \tag{D.23}$$

and

$$\mathbb{E}[(\alpha + k\delta(\eta_2))Q] = \frac{1}{2}\mathbb{E}[(2\alpha\delta(\eta_2) + x\delta'(\eta_2))] . \tag{D.24}$$

Let $P_{x|\eta_2}$ and P_{η_2} denote the conditional distribution of x given η_2 and the marginal distribution of η_2 respectively. We then get

$$\begin{aligned} \mathbb{E}[\delta(\eta_2)Q] &= \int\int x\operatorname{sgn}(\eta_2)dP_{x|\eta_2}\delta(\eta_2)dP_{\eta_2} \\ &= \int \mathbb{E}[x|\eta_2]\operatorname{sgn}(\eta_2)\delta(\eta_2)dP_{\eta_2} = 0 . \end{aligned} \tag{D.25}$$

Moreover

$$\begin{aligned}\mathbb{E}[x\delta'(\eta_2))] &= \int\int x\, dP_{x|\eta_2}\, \delta'(\eta_2)\, dP_{\eta_2} \\ &= \int \mathbb{E}[x|\eta_2]\, \delta'(\eta_2)\, \Phi_{\eta_2}\, d\eta_2 \\ &= -\mathbb{E}'[x|\eta_2 = 0]\, \Phi_{\eta_2}(0) - \mathbb{E}[x|\eta_2 = 0]\, \Phi'_{\eta_2}(0) \\ &= -\mathbb{E}'[x|\eta_2 = 0]\, \Phi_{\eta_2}(0) \qquad \text{(D.26)}\end{aligned}$$

where Φ_{η_2} is the (stationary) marginal density of η_2 and $'$ indicates the derivative with respect to η_2. Using (D.25) and (D.26) in (D.24) we obtain

$$\begin{aligned}\alpha\mathbb{E}[Q] &= \alpha\mathbb{E}[\delta(\eta_2)] + \frac{1}{2}\mathbb{E}[x\delta'(\eta_2)] \\ &= \alpha\Phi_{\eta_2}(0) - \frac{1}{2}\mathbb{E}'[x|\eta_2 = 0]\, \Phi_{\eta_2}(0) \\ &= k(\alpha - \frac{1}{2}\mathbb{E}'[x|\eta_2 = 0])\,.\end{aligned}$$

And with the following approximation[2]

$$\mathbb{E}'[x|\eta_2 = 0] \approx \frac{\mathbb{E}[x\eta_2]}{\mathrm{Var}[\eta_2]} = \frac{\mathbb{E}[F]}{\mathrm{Var}[\eta_2]}\,,$$

we obtain

$$\alpha\mathbb{E}[Q] \approx k(\alpha - k^2\mathbb{E}[F]). \qquad \text{(D.27)}$$

Solving (D.23) and (D.27) yields

$$\mathbb{E}[F] \approx \frac{\alpha}{\alpha + k^2}\,, \quad \text{and} \quad \mathbb{E}[Q] \approx k\mathbb{E}[F]\,.$$

The expectation of the integral appearing in the expression of $\mathbb{E}[Z(\eta_1, r)]$ is therefore

$$\begin{aligned}\mathbb{E}\Big(\int_0^t \alpha e^{\alpha(s-t)}\, (\eta_2(t) + \alpha k\, \mathrm{sgn}(\eta_2(t)))\, dW_s\Big) &\approx \alpha k\mathbb{E}[Q] + \mathbb{E}[F] \\ &\approx (\alpha k^2 + 1)\mathbb{E}[F] \\ &\approx \frac{\alpha(\alpha k^2 + 1)}{\alpha + k^2}\,. \qquad \text{(D.28)}\end{aligned}$$

References

[1] V. I. Utkin, J. Guldner, and J. Shi, *Sliding mode control in electromechanical systems*, ser. The Taylor & Francis systems and control book series. London, Philadelphia, PA: Taylor & Francis, 1999.

[2] Follows by approximating x with $x = a\eta_2 + \epsilon$ where ϵ is an uncorrelated function and a is a constant.

[2] J. Liu and X. Wang, *Advanced Sliding Mode Control for Mechanical Systems : Design, Analysis and MATLAB Simulation*. Berlin: Springer Berlin, 2012, 11,N15. [Online]. Available: http://opac.inria.fr/record=b1133965

[3] M. Herrera, W. Chamorro, A. P. Gomez, and O. Camacho, "Sliding mode control: An approach to control a quadrotor," *Computer Aided System Engineering (APCASE), 2015 Asia-Pacific Conference on*, pp. 314–319, 2015.

[4] N. Sakamoto, T. Niimura, K. Ozawa, and H. Takamori, "Robust feedback control for the subsidy policy about plug-in electric vehicle using sliding mode control," *Elect. Eng. Jpn.*, vol. 194, no. 1, pp. 10–17, 2016.

[5] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, ser. Applications of Mathematics (New York). Berlin: Springer-Verlag, 1992, vol. 23.

[6] B. Øksendal, *Stochastic differential equations*, 6th ed., ser. Universitext. Berlin: Springer-Verlag, 2003, an introduction with applications. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-14394-6

[7] X. Mao and C. Yuan, *Stochastic differential equations with Markovian switching*. London: Imperial College Press, 2006.

[8] A. K. Zvonkin, "A transformation of the phase space of a diffusion process that will remove the drift," *Mat. Sb. (N.S.)*, vol. 93(135), pp. 129–149, 152, 1974.

[9] A. J. Veretennikov, "On strong solutions and explicit formulas for solutions of stochastic integral equations," *Math. USSR Sb.*, vol. 39, no. 3, pp. 387–403, 1981. [Online]. Available: http://dx.doi.org/10.1070/SM1981v039n03ABEH001522

[10] D. Sundararajan, *The discrete Fourier transform*. World Scientific Publishing Co., Inc., River Edge, NJ, 2001, theory, algorithms and applications.

[11] H. Risken, *The Fokker-Planck equation*, 2nd ed., ser. Springer Series in Synergetics. Berlin: Springer-Verlag, 1989, vol. 18, methods of solution and applications. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-61544-3

[12] L. Arnold, *Stochastic differential equations: theory and applications*. New York: Wiley-Interscience [John Wiley & Sons], 1974, translated from the German.

Part III

Appendix

Details

This chapter presents further details and perspectives on parts of the thesis already discussed.

1 A Stochastic Model with Markovian Switching

Inspired by the model presented in paper C, the following is an extension of this model.

Let the stochastic process $x(t)$ be driven by the SDE

$$dx_t = f_q(x_t)dt + \beta dW_t, \quad q(t) \in Q = \{1,2\} \tag{1}$$

where β is a constant and

$$f_q(x_t) = \begin{cases} v_1 & \text{if } q = 1 \\ -v_2 & \text{if } q = 2 \end{cases}$$

with the constraints that $v_1, v_2 > 0$. The discrete process $q(t)$ is driven by a state-dependent Markov process where the transition probabilities are given according to

$$\begin{aligned} \lim_{h\to 0} P(q(t+h) = 1 | q(t) = 2 \wedge x(t) < 0) - \lambda h &= O(h^2) \quad \text{and} \\ \lim_{h\to 0} P(q(t+h) = 2 | q(t) = 1 \wedge x(t) > 0) - \mu h &= O(h^2) \end{aligned}$$

where λ and μ are the intensities of the Markov process. The state dependent switching activity is further outlined in Table 1. A solution to (1) can be observed as the two interconnected processes $(x(t), q(t))$.

The dynamics of the system causes two Fokker Planck equations defined on $\mathbb{R}$, one for each discrete dynamics. Let the corresponding "densities" be defined as $\phi_1(t,x)$ and $\phi_2(t,x)$. Then the Fokker Planck equation gives the relation

$$\frac{\partial \phi_1}{\partial t} = -v_1 \frac{\partial \phi_1}{\partial x} + \frac{\beta^2}{2} \frac{\partial^2 \phi_1}{\partial x^2} + \lambda \mathbb{I}_{(x<0)} \phi_2 - \mu \mathbb{I}_{(x>0)} \phi_1 \tag{2a}$$

$$\frac{\partial \phi_2}{\partial t} = v_2 \frac{\partial \phi_2}{\partial x} + \frac{\beta^2}{2} \frac{\partial^2 \phi_2}{\partial x^2} + \mu \mathbb{I}_{(x>0)} \phi_1 - \lambda \mathbb{I}_{(x<0)} \phi_2 \tag{2b}$$

	$q(t)$	switching	intensity
$x(t) < 0$	1	inactive	-
$x(t) > 0$	1	active	μ
$x(t) < 0$	2	active	λ
$x(t) > 0$	2	inactive	-

Table 1: An outline of the switching activity.

The indicator functions are included to ensure that ϕ_i, for $i \in \{1,2\}$, only affects the time-derivative in the corresponding domain. Furthermore, it is assumed that

$$\int_{\mathbb{R}} \phi_1(t,x) + \phi_2(t,x)dx \quad = \quad 1\,. \tag{3}$$

with the restriction that $\phi_1(t,x), \phi_2(t,x) \geq 0$ for all $t \geq 0, x \in \mathbb{R}$, such that the total probability mass assigned to the two functions is maintained even though the two functions lose or gain probability mass from each other. Furthermore, the dynamics of the Fokker-Planck equations is interconnected so their solutions share boundary conditions.

1.1 Application of Fourier Transformation

Application of the Fourier transformation to the terms in the Fokker-Planck equation in (2) gives for $i \in \{1,2\}$

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{ikx} \frac{\partial}{\partial t}\phi_i(t,x)dx &= \frac{\partial}{\partial t}\hat{\phi}_i(t,k) \\
\int_{-\infty}^{\infty} e^{ikx} \frac{\partial}{\partial x}\phi_i(t,x)dx &= -ik\hat{\phi}_i(t,k) \\
\int_{-\infty}^{\infty} e^{ikx} \frac{\partial^2}{\partial x^2}\phi_i(t,x)dx &= -k^2\hat{\phi}_i(t,k)\,.
\end{aligned}$$

Furthermore, an expression for the expectation $\mathbb{E}[x(t)]$ can be obtained by differentiating the Fourier transformed functions $\hat{\phi}_1(t,k) + \hat{\phi}_2(t,k)$ with respect to k and afterwards to set $k = 0$, that is,

$$E[x(t)] \quad = \quad -i\frac{\partial}{\partial k}\left(\hat{\phi}_1(t,k) + \hat{\phi}_2(t,k)\right)|_{k=0}\,.$$

By applying the same procedure to the sum of the Fourier transformed Fokker-Planck equations

$$\frac{\partial}{\partial t}\hat{\phi}_1(t,k) + \frac{\partial}{\partial t}\hat{\phi}_2(t,k) = ik\left(v_1\hat{\phi}_1(t,k) - v_2\hat{\phi}_2(t,k)\right) - k^2\frac{\beta^2}{2}\left(\hat{\phi}_1(t,k) + \hat{\phi}_2(t,k)\right)\,,$$

observe that

$$\frac{\partial}{\partial k}\left(\frac{\partial}{\partial t}\hat{\phi}_1(t,k)+\frac{\partial}{\partial t}\hat{\phi}_2(t,k)\right)|_{k=0} \quad = \quad i\left(v_1\hat{\phi}_1(t,0)-v_2\hat{\phi}_2(t,0)\right) \ .$$

Hence, by interchanging the differential order, and multiplying with the imaginary unit i gives the relation

$$\frac{\partial}{\partial t}E[x(t)] \quad = \quad v_1\hat{\phi}_1(t,0)-v_2\hat{\phi}_2(t,0) \tag{4}$$

which has the solution

$$E[x(t)] \quad = \quad E[x(0)]+\int_0^t v_1\hat{\phi}_1(\tau,0)-v_2\hat{\phi}_2(\tau,0)d\tau \tag{5}$$

Due to (3) it follows that,

$$\begin{aligned} 1 \quad &= \quad \int_{-\infty}^{\infty}\phi_1(t,x)+\phi_2(t,x)dx \\ &= \quad \int_{-\infty}^{\infty}e^{i\cdot 0\cdot x}\left(\phi_1(t,x)+\phi_2(t,x)\right)dx \quad = \quad \hat{\phi}_1(t,0)+\hat{\phi}_2(t,0) \end{aligned}$$

so $\hat{\phi}_2(t,0)=1-\hat{\phi}_1(t,0)$. By substituting this into (5) gives the expression

$$\begin{aligned} E[x(t)] \quad &= \quad E[x(0)]+\int_0^t\left(-v_2+(v_1+v_2)\hat{\phi}_1(t,0)\right)d\tau \\ &= \quad E[x(0)]-v_2t+(v_1+v_2)\int_0^t\hat{\phi}_1(t,0)d\tau \end{aligned}$$

This result indicates that the expectation of the process $x(t)$ diverges as $t\to\infty$.

SUMMARY

This thesis treats stochastic systems with switching dynamics. Models with these characteristics are studied from several perspectives. Initially in a simple framework given in the form of stochastic differential equations and, later, in an extended form which fits into the framework of sliding mode control. It is investigated how to understand and interpret solutions to models of switched systems, which are exposed to discontinuous dynamics and uncertainties (primarily) in the form of white noise. The goal is to gain knowledge about the performance of the system by interpreting the solution and/or its probabilistic properties.

ISSN (online): 2446-1628
ISBN (online): 978-87-7112-970-0

AALBORG **UNIVERSITETSFORLAG**